matics Library

JUN 18 1947

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

R. BRAUER
UNIVERSITY OF TORONTO

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

L. M. GRAVES
UNIVERSITY OF CHICAGO

H. WHITNEY
HARVARD UNIVERSITY

A. WINTNER
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

S. BOCHNER	G. BIRKHOFF	L. BERS
D. MONTGOMERY	R. COURANT	J. L. DOOB
C. L. SIEGEL	P. HARTMAN	N. JACOBSON
D. C. SPENCER	J. L. VANDERSLICE	D. C. LEWIS
M. A. ZORN	A. ZYGMUND	A. D. WALLACE

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXIX, Number 2
APRIL, 1947

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

CONTENTS

	PAGE
On the ergodic theorems. By PHILIP HARTMAN,	193
Mapping theorems for non-compact spaces. By C. H. DOWKER,	200
On the zeta-functions of algebraic number fields. By RICHARD BRAUER,	243
Asymptotic integrations of the adiabatic oscillator. By AUREL WINTNER,	251
Polynomial least square approximations. By D. C. LEWIS,	273
A note on Laplace transforms of functions whose spectra are confined to a given set. By E. K. HAVILAND,	279
On the representations of the symmetric group (second paper). By G. DE B. ROBINSON,	286
Polynomial matrices in one variable, differential equations and module theory. By ERNST SNAPPER,	299
Systems of total differential equations and Liouville's theorem on conformal mappings. By PHILIP HARTMAN,	327
Generic algebras. By W. J. R. CROSBY,	333
Simple links and fixed sets under continuous mappings. By J. L. KELLEY,	348
Inequalities of higher degree in one unknown. By BRUCE ELWYN MESERVE,	357
On Riesz summability and summability by Dirichlet's series. By C. T. RAJAGOPAL,	371
A study on the theory of conjugate nets. By CHUAN-CHIH HSIUNG,	379
Six-rings in minimal five-color maps. By ARTHUR BERNHART,	391
A note on finitely-additive measures. By ELLEN F. BUCK and R. C. BUCK,	413

The AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications may be sent to Professor F. D. MURNAGHAN at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE 18, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

ON THE ERGODIC THEOREMS.*

By PHILIP HARTMAN.

It was recently suggested to me by Professor Wintner that it should be possible to find a proof of the Birkhoff ergodic theorem [1] along the lines of the simplest proof of the fundamental theorem of calculus, namely, a proof depending on a lemma of F. Riesz [4]. The role of Riesz's lemma in the Lebesgue theory of differentiation is the avoidance of Vitali's covering theorem on which the earlier proofs of the fundamental theorem of calculus were based. In the proof of the ergodic theorem, as given by Birkhoff and his successors, the corresponding role of Vitali's covering theorem is represented by Birkhoff's decomposition of the image sets or by other covering theorems. As will be seen in the proof of (I), below, an appeal to the Riesz lemma leads easily to an improved form of the fundamental inequality on which Birkhoff's proof of the ergodic theorem depends. This improved inequality was also obtained by Kakutani and Yosida [3] using the methods of Birkhoff.

For the sake of completeness, the Birkhoff ergodic theorem will be deduced from this inequality in (II) and the (L^p), $p \geqq 1$, mean ergodic theorem will then be derived from (I) and (II). This procedure is precisely the reverse of that applied by Wiener [5] and his followers.

For earlier references, see Hopf [2].

In what follows, Ω denotes a space of points P carrying a non-negative measure μ, for which $\mu(\Omega) = 1$. Let $P_t = \tau_t P$, where $-\infty < t < \infty$, be a group of measure-preserving transformations of Ω into itself; by this is meant that

$$\tau_u(\tau_v P) = \tau_{u+v} P \tag{1}$$

and each transformation τ_t sends measurable sets into measurable sets of the same measure. Let $f(P)$ be a real-valued function of class $(L^1) = (L)$ on Ω, that is, $f(P)$ is measurable and

$$\int_\Omega |f(P)|\, d\mu < \infty. \tag{2}$$

In addition, it will be supposed that $f(\tau_t P)$ is a measurable function on the

* Received February 19, 1947.

product space of Ω and the line $-\infty < t < \infty$, where measure on the line is the ordinary Lebesgue measure. If these assumptions are satisfied and if a zero set of points P is excluded, the integral

$$\int_0^t f(\tau_u P)\,du$$

exists for all finite values of t by virtue of the measure-preserving character of τ_t and by Fubini's theorem on iterated integrals.

A subset S of Ω will be said to be invariant if it is measurable and if its image $\tau_t S$, under each transformation τ_t, differs from S only by a zero set, which may depend on t.

(I) *Let α be an arbitrary number and let U denote the set of points P of Ω for which*

$$\text{l. u. b.}_{0<t<\infty}\ t^{-1}\int_0^t f(\tau_u P)\,du > \alpha.$$

Then, if S is any invariant subset of Ω,

$$\alpha\mu(S\cdot U) \leqq \int_{S\cdot U} f(P)\,d\mu. \tag{3}$$

In particular, if S is contained in U,

$$\alpha\mu(S) \leqq \int_S f(P)\,d\mu. \tag{3 bis}$$

The form of Riesz's lemma [4] which will be needed for the proof of (I) is as follows:

Let $g(x)$ be a real-valued continuous function on an interval $a \leqq x \leqq b$ and let α be any fixed number. Let E denote the set of interior points x of (a, b) for which there exists at least one number x' satisfying $a \leqq x' < x$ and $g(x) - g(x') > \alpha(x - x')$. Then E is an open set and the inequality $\alpha(b_n - a_n) \leqq g(b_n) - g(a_n)$ holds for every n if $\Sigma(a_n, b_n)$ is the decomposition of E into disjoint open intervals.

Proof of (I). For a fixed t, let $U(t)$ denote the set of points P of Ω for which

$$\text{l. u. b.}_{0<v<t}\ v^{-1}\int_0^v f(\tau_u P)\,du > \alpha.$$

Then $U(t)$ is a measurable set which is non-decreasing with t and its limit set is U,

$$U = \lim_{t\to\infty} U(t). \tag{4}$$

Let $S(t)$ denote the image of $U(t)$ under the transformation τ_t, so that $S(t) = \tau_t U(t)$ or $U(t) = \tau_{-t} S(t)$. By (1),

$$\int_{t-v}^{t} f(\tau_{-u}P)\,du = \int_0^v f(\tau_u \tau_{-t} P)\,du, \tag{5}$$

hence $S(t)$ is the set of points P for which

$$\underset{0<v<t}{\text{l. u. b.}}\ v^{-1}\int_{t-v}^{t} f(\tau_{-u}P)\,du > \alpha. \tag{6}$$

For a point P and a positive number m, let $E(P, m)$ denote the set of t-values on $(0, m)$ for which (6) holds. If $g(x)$ is defined to be

$$\int_0^x f(\tau_{-u}P)\,du, \qquad (0 \leqq x \leqq m),$$

then the set $E(P, m)$ has the same significance as the set E in Riesz's lemma. Hence, on adding the inequalities implied by this lemma,

$$\alpha \Sigma (b_n - a_n) \leqq \Sigma (g(b_n) - g(a_n))$$

or

$$\alpha \int_E dt \leqq \int_E f(\tau_{-t}P)\,dt, \qquad (E = E(P, m)).$$

Integrating this inequality over the measurable set S, one obtains

$$\alpha \int_S \Big(\int_E dt\Big)\,d\mu \leqq \int_S \Big(\int_E f(\tau_{-t}P)\,dt\Big)\,d\mu.$$

Hence, by Fubini's theorem on iterated integrals,

$$\alpha \int_0^m \mu(S \cdot S(t))\,dt \leqq \int_0^m \Big(\int_{S\cdot S(t)} f(\tau_{-t}P)\,d\mu\Big)\,dt.$$

Since S is an invariant set, it follows that, when t is fixed, $\tau_{-t}(S \cdot S(t)) = S \cdot \tau_{-t} S(t) = S \cdot U(t)$ if a zero set of points P of Ω is ignored. Consequently,

$$\int_{S\cdot S(t)} f(\tau_{-t}P)\,d\mu = \int_{S\cdot U(t)} f(P)\,d\mu$$

and

$$\mu(S \cdot S(t)) = \mu(S \cdot U(t))$$

since τ_{-t} is measure-preserving. Consequently, the last inequality can be written in the form

$$\alpha \int_0^m \mu(S \cdot U(t))\,dt \leqq \int_0^m \int_{S \cdot U(t)} f(P)\,d\mu\,dt. \tag{7}$$

On the other hand, the limit relation (4) implies that

$$\mu(S \cdot U(t)) \to \mu(S \cdot U) \text{ and } \int_{S \cdot U(t)} f(P)\,d\mu \to \int_{S \cdot U} f(P)\,d\mu \qquad (t \to \infty).$$

Therefore, if the inequality (7) is divided by m, the assertion (3) follows by letting $m \to \infty$. This completes the proof of (I).

An application of (I) to the function $-f(P)$ and to the constant $\alpha = -\beta$ gives

(I bis) *Let β be an arbitrary number and let V denote the set of points P of Ω for which*

$$\underset{0<t<\infty}{\text{g. l. b.}}\ t^{-1} \int_0^t f(\tau_u P)\,du < \beta. \tag{8}$$

Then, if S is any invariant subset of Ω,

$$\beta\mu(S \cdot V) \geqq \int_{S \cdot V} f(P)\,d\mu; \tag{9}$$

and, if S is contained in V,

$$\beta\mu(S) \geqq \int_S f(P)\,d\mu. \tag{9 bis}$$

The statements (I) and (I bis) imply the following theorem, which contains the Birkhoff ergodic theorem:

(II) *For almost all points P of Ω,*

$$f^*(P) = \lim_{t \to \infty} t^{-1} \int_0^t f(\tau_u P)\,du \tag{10}$$

exists, is a function of class (L), and

$$\int_\Omega f^*(P)\,d\mu = \int_\Omega f(P)\,d\mu. \tag{11}$$

Proof of (II). Let α and β be any pair of numbers satisfying $\alpha > \beta$ and let S denote the set of points P for which

$$\liminf_{t \to \infty} t^{-1} \int_0^t f(\tau_u P)\,du < \beta < \alpha < \limsup_{t \to \infty} t^{-1} \int_0^t f(\tau_u P)\,du.$$

Then, in virtue of (1), S is an invariant set. Also, S is contained in the sets

U and V, defined in (I) and (I bis), respectively. Hence, (I) and (I bis) imply that

$$\alpha\mu(S) \leqq \int_S f(P)\,d\mu \leqq \beta\mu(S).$$

Since $\alpha > \beta$ and $\mu(S) \leqq \mu(\Omega) = 1$, it follows that $\mu(S) = 0$. Hence, the limit (10) exists for almost all P; where, however, the possibility that this limit is $\pm\infty$ is not yet excluded.

In order to see that $f^*(P)$ is finite for almost all P, let S denote the set of points P for which $f^*(P) = +\infty$. Then S is an invariant subset of Ω and S is contained in the set U of (I) for every constant α. Thus (3 bis) holds for every α. Hence, by (2), $\mu(S) = 0$. It is similarly shown that the set of points P for which $f^*(P) = -\infty$ is a zero set.

Finally, to prove that $f^*(P)$ is of class (L), let α and β be any pair of numbers such that $\alpha < \beta$ and let S be the set of points P for which $f^*(P)$ exists and satisfies

$$\alpha < f^*(P) < \beta.$$

Then, (I) and (I bis) imply that

$$\alpha\mu(S) \leqq \int_S f(P)\,d\mu \leqq \beta\mu(S).$$

It follows, from the definition of Lebesgue integrals in terms of upper and lower sums connected with a mesh on the f^*-axis, that $f^*(P)$ is of class (L) and that its integral satisfies (11).

This completes the proof of (II).

(III) *If the function $f(P)$ in* (II) *is of class (L^p), $p \geqq 1$, on Ω then so is the function $f^*(P)$, defined by* (10), *and*

$$\int_\Omega \left| f^*(P) - t^{-1}\int_0^t f(\tau_u P)\,du \right|^p d\mu \to 0 \text{ as } t \to \infty. \tag{12}$$

Proof of (III). By Hölder's inequality,

$$\left| t^{-1}\int_0^t f(\tau_u P)\,du \right|^p \leqq t^{-1}\int_0^t \left| f(\tau_u P) \right|^p du. \tag{13}$$

The arguments used before the statement of (I) show that, for almost all P, the integral on the right of (13) exists for all finite t and is of class (L) for every fixed t. Hence, the integral in (12) exists if $f^*(P)$ is of class (L^p). But, by (10) and (13),

$$\left| f^*(P) \right|^p \leqq \lim_{t\to\infty} t^{-1}\int_0^t \left| f(\tau_u P) \right|^p du,$$

where the limit on the right exists for almost all P and is of class (L) by virtue of an application of (II) to the function $|f(P)|^p$. Consequently, $f^*(P)$ if of class (L^p).

The assertion (10) of (II) is equivalent to

$$|f^*(P) - t^{-1}\int_0^t f(\tau_u P)\,du|^p \to 0, \qquad (t \to \infty), \tag{14}$$

while the assertion of (III) is that (14) can be integrated term-by-term. On the other hand, this term-by-term integration is allowed if

$$\int_T |f^*(P) - t^{-1}\int_0^t f(\tau_u P)\,du|^p\,d\mu \to 0 \text{ as } \alpha \to \infty \tag{15}$$

holds uniformly in t, where $T = T_\alpha$ is the set of points P for which

$$\underset{0<t<\infty}{\text{l. u. b.}}\ |f^*(P) - t^{-1}\int_0^t f(\tau_u P)\,du| > \alpha. \tag{16}$$

By the Minkowski inequality, the integral in (15) is majorized by the two integrals

$$\int_T |f^*(P)|^p\,d\mu \tag{17}$$

and

$$\int_T |t^{-1}\int_0^t f(\tau_u P)\,du|^p\,d\mu.$$

But the Hölder inequality (13) shows that the last integral does not exceed

$$\int_T (t^{-1}\int_0^t |f(\tau_u P)|^p\,du)\,d\mu,$$

which, by Fubini's theorem on iterated integrals, has the value

$$t^{-1}\int_0^t (\int_T |f(\tau_u P)|^p\,d\mu)\,du$$

and is, therefore, majorized by

$$\underset{0<u<\infty}{\text{l. u. b.}} \int_W |f(P)|^p\,d\mu, \qquad (W = \tau_u T). \tag{18}$$

In virtue of the absolute continuity of the integrals

$$\int |f^*(P)|^p\,d\mu \text{ and } \int |f(P)|^p\,d\mu,$$

and the measure-preserving character of the transformations τ_u, the quantities (17) and (18) tend to 0 as $\alpha \to \infty$ if the measure of $T = T_\alpha$ does.

By (1), the function (10) is invariant, that is $f^*(\tau_t P) = f^*(P)$ for every t. Hence,

$$| f^*(P) - t^{-1} \int_0^t f(\tau_u P)\, du | \leqq t^{-1} \int_0^t | f^*(\tau_u P) - f(\tau_u P) |\, du,$$

and so the set $U = U_\alpha$ of points P for which

$$\underset{0<t<\infty}{\text{l. u. b.}}\ t^{-1} \int_0^t | f^*(\tau_u P) - f(\tau_u P) |\, du > \alpha$$

contains the set $T = T_\alpha$. If, in (I), the function $f(P)$ is replaced by $| f^*(P) - f(P) |$ and the invariant set S is chosen to be the entire space Ω, (3) gives

$$\alpha\mu(U) \leqq \int_U | f^*(P) - f(P) |\, d\mu.$$

Thus,

$$\mu(T) \leqq \mu(U) \leqq \alpha^{-1} \int_\Omega | f^*(P) - f(P) |\, d\mu \to 0, \text{ as } \alpha \to \infty.$$

This completes the proof of (III).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] G. D. Birkhoff, "Proof of the ergodic theorem," *Proceedings of the National Academy of Sciences*, vol. 18 (1931), pp. 656-660.

[2] E. Hopf, *Ergodentheorie*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 5 (1937), no. 2.

[3] S. Kakutani and K. Yosida, "Birkhoff's ergodic theorem and the maximal ergodic theorem," *Proceedings of the Imperial Academy of Tokyo*, vol. 15 (1939), pp. 165-168.

[4] F. Riesz, "Sur l'existence de la dérivée des fonctions monotones et sur quelques problèmes qui s'y rattachent," *Acta Litterarium ac Scientiarum* (Szeged), vol. 5 (1932), pp. 208-221.

[5] N. Wiener, "The ergodic theorem," *Duke Mathematical Journal*, vol. 5 (1939), pp. 1-18.

MAPPING THEORMS FOR NON-COMPACT SPACES.*

By C. H. Dowker.

The chief aim of this paper is to extend to more general spaces the theorems of Hopf [17] [1] and Bruschlinsky [7] on the mappings of complexes in spheres. Since these theorems can be stated most neatly in terms of cohomologies, the latter are used throughout the paper.

It is shown that, if cocycles based on infinite coverings are used, the theorems of Hopf and Bruschlinsky can be extended to all paracompact [2] normal spaces, a class of spaces which includes in particular the compact [3] Hausdorff spaces and the metric separable spaces. With the traditional Čech cocycles based on finite coverings, the theorems hold for countably compact normal spaces. If the cocycles are based on the finite coverings of compact subsets, Brushlinsky's theorem, but not Hopf's, extends to a certain class [4] of non-compact spaces. Keeping cocycles based on finite coverings but replacing homotopy by uniform homotopy in the statements of Hopf's and Bruschlinsky's theorems, one can extend these theorems to arbitrary normal spaces.

We call attention to the following additional results: (a) The 1-dimensional Čech co-Betti number of the straight line is the power of the continuum. (b) The covering dimension of a normal space is the same whether based on finite coverings or on more general locally finite coverings. (c) Borsuk's theorem ([5], p. 103) on extensions of uniformly homotopic mappings into absolute neighborhood retracts holds for normal spaces, and the corresponding theorem with ordinary homotopy holds for paracompact normal spaces.

1. Coverings and nerves. By a covering of a space [5] we mean a

* Received May 31, 1940; Revised October 24, 1946. The author wishes to thank E. G. Begle for his assistance in an extensive revision of this paper. Some of the results in the paper were reported in the *Proceedings of the National Academy of Sciences*, vol. 23 (1937), pp. 293-294. Section **3** was presented to the American Mathematical Society Febuary 23, 1946, under the title of "Lebesgue dimension of a normal space."

[1] Numbers in square brackets refer to the bibliography.

[2] Paracompact spaces have been investigated by J. Dieudonné [11].

[3] Compact means bicompact.

[4] This is the class of locally connected locally compact paracompact normal spaces.

[5] In general no separation axiom is assured. Thus, using the nomenclature of Alexandroff-Hopf [3], a space is required to be a topological space but not necessarily a T_0 space.

covering by a finite or infinite collection of open sets. By the nerve of a covering we mean the nerve realized as a space with the Euclidean metric as defined by S. Lefschetz ([23], p. 5). Each point p of such a Euclidean complex has a set of barycentric coordinates $\{x_i\}$, a coordinate x_i being associated with each vertex a_i of the complex, such that $\Sigma x_i = 1$, and such that $0 \leq x_i \leq 1$ for every i, $x_i > 0$ if and only if a_i is a vertex of the simplex containing p, and $x_i = 1$ if and only if p is the vertex a_i. The distance between any two points $\{x_i\}$ and $\{y_i\}$ of the complex is given by the formula $(\Sigma \mid x_i - y_i \mid^2)^{\frac{1}{2}}$.

A covering of a space is called *finite* if it consists of a finite number of open sets. A space is called compact if every covering of the space has a finite refinement.[6] Čech's homology theory [8] is based on finite coverings and has proved very useful for investigating compact spaces. For non-compact spaces, even those as simple as the straight line, this theory gives intuitively unsatisfactory results as is shown in Theorem 9. 6. As might be supposed, the coverings appropriate for use with non-compact spaces are infinite. We shall see that important theorems can be extended to a broad class of non-compact spaces if the finiteness conditions are relaxed.

A covering of a space is called *locally finite*[7] if each point of R has a neighborhood meeting only a finite number of sets of the covering. A space is called *paracompact* if every covering of the space has a locally finite refinement. Paracompact spaces are thus a generaliaztion of compact spaces. J. Dieudonné [11] has shown that many properties of compact spaces extend to paracompact spaces. In particular, he has shown that for paracompact spaces normality is less restrictive than the Hausdorff separation axiom. In the theorems of this paper we require that paracompact spaces be normal but not necessarily Hausdorff.

A covering of a space is called *star-finite* if each set of the covering meets only a finite number of other sets of the covering. A space is called an *s-space* if every covering has a star-finite refinement. The star-finite coverings include the finite coverings and are included in the locally finite coverings. They are from one point of view the most immediate generalization of finite coverings, for the star-finite coverings have as nerves locally finite complexes, and locally finite complexes are the simplest type after the finite complexes. In the same

---

[6] A covering $\mathfrak{U}$ of R is a refinement of a covering $\mathfrak{V}$ if every open set of $\mathfrak{U}$ is contained in some open set of $\mathfrak{V}$.

[7] Locally finite = neighborhood-finite. "Locally finite" has been used by S. Lefschetz ([24], p. 13) to mean star-finite, and by A. Weil ([26], p. 34) to mean that each compact set meets a finite number of open sets of the covering.

sense the s-spaces are the most immediate generalization of the compact spaces. The generalization is moreover non-trivial since, as has been shown by S. Kaplan [21], every separable metric space is an s-space. The paracompact spaces are a further non-trivial generalization. An infinite complex is paracompact [8] but it need not be an s-space.

In combinatorial topology the chief interest in locally finite coverings results from the fact that the theorem of Alexandroff ([1], p. 121) [9] on mapping a space into the nerve of a covering can be extended to normal spaces if the coverings are required to be locally finite. Theorem 1.1, which is a form of Alexandroff's theorem, is the basic lemma of this paper, being used in the proof of all the remaining theorems.

Let $\mathfrak{U}$ be any covering of a space R and let N be the nerve of $\mathfrak{U}$. A mapping ϕ of R into N is called *canonical* relative to $\mathfrak{U}$ if the inverse image of the star of each vertex of N is contained in the open set of $\mathfrak{U}$ corresponding to this vertex. Let $\{U_i\}$, where each i is a transfinite ordinal, be the open sets of $\mathfrak{U}$ and let $\{u_i\}$ be the corresponding vertices of N. Let $St(u_i)$ be the star of u_i in N. Then ϕ is canonical when $\phi^{-1}St(u_i) \subset U_i$, for each $U_i \in \mathfrak{U}$.

If $\mathfrak{U}$ is a point-finite covering, i. e., if each point of R is in only a finite number of sets of $\mathfrak{U}$, there is another useful formulation of the condition that a mapping of R in the nerve of $\mathfrak{U}$ be canonical. Let p be any point of R. The vertices of N corresponding to those open sets of the point-finite covering $\mathfrak{U}$ which contain p are the vertices of a certain simplex $\sigma(p)$ of N. This simplex $\sigma(p)$ is called the simplex of N determined by p. *A necessary and sufficient condition* that a mapping ϕ of R in the nerve N of a point finite covering $\mathfrak{U}$ be canonical is that for every point $p \in R$, $\phi(p)$ is in the closure $\overline{\sigma(p)}$ of the simplex of N determined by p. For, if ϕ is canonical, let p be any point of R. Then, since $\phi(R) \subset N$, $\phi(p) \in \sigma$ for some simplex σ of N. Let u_i be any vertex of σ. Then $\phi(p) \in St(u_i)$, and hence $p \in \phi^{-1}St(u_i)$. Hence $p \in U_i$ and u_i is a vertex of $\sigma(p)$. Hence all vertices of σ are vertices of $\sigma(p)$ and $\sigma \subset \overline{\sigma(p)}$. Hence $\phi(p) \in \overline{\sigma(p)}$. On the other hand, let $\phi(p) \in \overline{\sigma(p)}$ for each point $p \in R$. Let U_i be an arbitrary open set of $\mathfrak{U}$ and let p be an arbitrary point of $\phi^{-1}St(u_i)$. Then $\phi(p) \in St(u_i)$ and also $\phi(p) \in \overline{\sigma(p)}$. Hence u_i is a vertex of $\sigma(p)$. Hence $p \in U_i$. Hence $\phi^{-1}St(u_i) \subset U_i$. Therefore ϕ is canonical.

THEOREM 1.1. *If $\mathfrak{U}$ is a locally finite covering of a normal space R there is a canonical mapping ϕ of R into the nerve N of $\mathfrak{U}$.*

---

[8] The proof of this statement will be published in a forthcoming paper by the author.

[9] See also [22] and [18]. The extension to star-finite coverings of normal spaces is given by S. Lefschetz ([25], p. 41).

Proof. Let $\{U_i\}$ $(i = 1, 2, \cdots, \alpha$, where α is a finite or transfinite ordinal number) be the open sets of $\mathfrak{U}$ and let $\{u_i\}$ be the corresponding vertices of K. Let $F_{ijk\ldots m} = R - \Sigma_{r \neq i,j,\ldots,m} U_r$ be defined for every set of subscripts $(i, j, \cdots, m)$ such that $U_i U_j \cdots U_m \neq 0$, i. e., such that $u_i u_j \cdots u_m$ is a simplex of N. Any F may of course be the null set, but each point of R is contained in some F. Each F is closed.

A mapping ϕ of R into N will be defined by assigning to each point of R the barycentric coordinates of the corresponding point of N. A point of the closed simplex $u_i u_j \cdots u_m$ of N has coordinates $x_i^{(ij\ldots m)}$, $x_j^{(ij\ldots m)}, \cdots,$ $x_m^{(ij\ldots m)}$ where $x_i^{(ijk\ldots m)} = x_i^{(jik\ldots m)}$, etc. If the point is on the boundary simplex $u_j \cdots u_m$, its coordinates satisfy the relations $x_j^{(ij\ldots m)} = x_j^{(j\ldots m)}$, etc.

Let $x_i^{(ij)} = x_i^{(ij)}(p)$ be a continuous function of p for p in F_{ij} such that $0 \leq x_i^{(ij)}(p) \leq 1$, $x_i^{(ij)}(p) = 0$ if $p \in F_j$ and $x_i^{(ij)} = 1$ if $p \in F_i$. The existence of such a function is proved by Urysohn's lemma. Let $x_j^{(ij)} = 1 - x_i^{(ij)}$. In this way we determine the barycentric coordinates of the image of each F_{ij}. We proceed by recursion. Assume that the coordinates of each $F_{ijk\ldots m}$ having q subscripts have been determined. Let $F_{ijk\ldots mn}$ have $q + 1$ subscripts. Let

$$\begin{aligned} y_i^{(ijk\ldots mn)} &= x_i^{(ik\ldots mn)} && \text{for} \quad p \in F_{ik\ldots mn}, \\ &\quad \cdots \\ &= x_i^{(ijk\ldots m)} && \text{for} \quad p \in F_{ijk\ldots m}, \\ &= 0 && \text{for} \quad p \in F_{jk\ldots mn}. \end{aligned}$$

It is thus defined as a continuous function of p for p in a certain closed set of $F_{ijk\ldots mn}$, and hence the function can be extended to a continuous function on the whole of $F_{ijk\ldots mn}$. Similarly $y_j^{(ijk\ldots m)}, \cdots, y_m^{(ijk\ldots mn)}$ are defined as functions of p for $p \in F_{ijk\ldots mn}$. Let

$$\begin{aligned} y_n^{(ijk\ldots mn)} &= x_n^{(jk\ldots mn)} && \text{for} \quad p \in F_{jk\ldots mn}, \\ &\quad \text{etc.} \\ &= 0 && \text{for} \quad p \in F_{ijk\ldots m}, \\ &= 1 \end{aligned}$$

for p in the closed set where $y_i^{(ijk\ldots mn)}(p) = y_j^{(ijk\ldots mn)}(p) = \cdots = y_m^{(ijk\ldots mn)} = 0$. Let

$$x_i^{(ijk\ldots mn)} = \frac{y_i^{(ij\ldots n)}}{y_i^{(ij\ldots n)} + y_j^{(ij\ldots n)} + \cdots + y_n^{(ij\ldots n)}},$$

etc. Thus the mapping is defined for $F_{ijk\ldots mn}$ and reduces on the "boundary" sets $F_{ijk\ldots m}$ etc. to that previously defined. Thus the mapping can be extended to all sets $F_{ijk\ldots mn}$ with $q+1$ subscripts. Proceeding *ad inf.*, the mapping can be extended to a mapping ϕ of the whole space R into N.

The mapping ϕ is continuous on each closed set $F_{ij\ldots}$ and hence on the sum of any finite number of these. Since the covering $\mathfrak{U}$ is locally finite each point of R has a neighborhood containing points of only a finite number of the U's and hence contained in a finite collection of the F's. Therefore ϕ is continuous in the neighborhood of each point of R and hence is a continuous mapping.

Let the simplex of N determined by p be $\sigma(p) = u_i u_j \cdots u_m$. Then p is not in U_r for $r \neq i, j, \cdots, m$. Hence $p \in F_{ij\ldots m}$. But ϕ maps $F_{ij\ldots m}$ into the closure of $u_i u_j \cdots u_m$. Hence $\phi(p) \in \overline{\sigma(p)}$. Hence ϕ is canonical. This completes the proof.

It may be noted that a simplicial mapping of one infinite complex in another is not necessarily continuous. Thus a mapping of an infinite complex K may be discontinous even if each finite subcomplex of K is mapped continuously.

LEMMA 1.2. *Let $\mathfrak{U}$ be a locally finite covering of a normal space R and let ϕ be a canonical mapping of R in the nerve N of $\mathfrak{U}$. Let N be mapped into a space S by a mapping f which is continuous on each finite subcomplex of N. Then the mapping $f\phi$ of R into S is continuous.*

Proof. Let p be any point of R. Since $\mathfrak{U}$ is locally finite the point p has a neighborhood W which intersects only a finite number of the open sets U_i of $\mathfrak{U}$. The corresponding vertices u_i of N determine a finite subcomplex of N which contains $\phi(W)$. This finite subcomplex is mapped continuously by f. Thus $f\phi$ maps W continuously into S so that $f\phi$ is continuous in the neighborhood of each point p of R and hence is continuous.

2. Homotopy. Let R and S be two spaces and let I be the closed line segment $[0, 1]$. Two mapings, f_0 and f_1, of R into S are said to be homotopic if there is a mapping h of $R \times I$ into S such that, for each $x \in R$, $h(x, 0) = f_0(x)$ and $h(x, 1) = f_1(x)$. The mapping h is called a homotopy of f_0 and f_1. The homotopy is said to be a uniform homotopy if S is a metric space and if for every $\epsilon > 0$ there is a $\delta > 0$ such that if $| t - t' | < \delta$ then the distance $\rho(h(x, t), h(x, t')) < \epsilon$ for all x in R. In this case f_0 is said to be uniformly homotopic to f_1.

In general the property of being a uniform homotopy depends on the metric of S. However, if S is a compact metric space the property is a topological one. For, if ϕ is a homeomorphism of S onto a space S_1, then ϕ is uniformly continuous, and hence if $\epsilon > 0$ there is a number $\eta > 0$ such that if the distance between two points of S is less than η their images by ϕ are less than ϵ apart. Let h be a uniform homotopy of two mappings of R into S. Let $\delta > 0$ be chosen so that if $|t - t'| < \delta$ then $\rho(h(x,t), h(x,t')) < \eta$. Then $\rho(\phi h(x,t), \phi h(x,t')) < \epsilon$ and hence ϕh is a uniform homotopy.

Both homotopy and uniform homotopy are equivalence relations, so the set of all mappings of R into S can be uniquely divided into classes of homotopic or uniformly homotopic mappings.

Borsuk's theorem ([5], p. 103) on the extensions of uniformly homotopic mappings can be extended to normal spaces as follows:

THEOREM 2.1. *(Borsuk's theorem). Let R' be a closed subset of a normal space R and let f_0 and f_1 be two uniformly homotopic mappings of R' into an absolute neighbrohood retract* [10] *S. If f_0 can be extended to a mapping F_0 of R into S then f_1 can be similarly extended to F_1 with F_0 and F_1 uniformly homotopic.*

Proof. We assume that S is a subset of the Hilbert cube I_ω. Let η be a number such that the closed η-neighborhood $\bar{U}$ of S, i. e., the set of all points of I_ω whose distance from S is less than or equal to η, can be retracted onto S.

It follows immediately from the definition of uniform homotopy that there is a finite number of mappings $h(x,0), h(x,\delta), h(x,2\delta), \cdots$ beginning with f_0 and ending with f_1 such that the distance between any two successive mappings is less than η. It is therefore sufficient to prove the theorem under the assumption that $\rho(f_0(x), f_1(x)) < \eta$ for every x in R'.

Since R is normal f_1 can be extended to a mapping G of R into I_ω. Let V be the set of points $x \in R$ such that $\rho(F_0(x), G(x)) < \eta$. Then V is an open set and, since we assume $\rho(f_0(x), f_1(x)) < \eta$ for $x \in R'$, $V \supset R'$. Let $s(x)$ be a continuous function of x, for $x \in R$, such that $0 \leq s(x) \leq 1$, $s(x) = 0$ if $x \in R - V$, and $s(x) = 1$ if $x \in R'$.

Let h_1 be the mapping of $R \times I$ into I_ω which maps (x,t) in the point dividing the segment $[F_0(x), G(x)]$ in the ratio $t : 1-t$. Then h_1 is a uniform homotopy. Also h_1 maps $V \times I$ into $\bar{U}$ for, if $x \in V$, $F_0(x) \in S$ and $\rho(F_0(x), h_1(x,t)) \leq \rho(F_0(x), G(x)) < \eta$. Let h_2 be the mapping of $R \times I$

[10] The definition of absolute neighborhood retract is given by K. Borsuk ([4], p. 222).

into $\bar{U}$ defined by $h_2(x,t) = h_1(x, t \cdot s(x))$. Then h_2 is a uniform homotopy, and, if $x \in R'$, $s(x) = 1$ and $h_2(x,1) = h_1(x,1) = G(x) = f_1(x)$. Let ϕ be a retraction of $\bar{U}$ onto S, and let h_3 be the mapping of $R \times I$ into S defined by $h_3(x,t) = \phi h_2(x,t)$. Since $\bar{U}$ is a closed, and hence compact, subset of I_ω, therefore ϕ is uniformly continuous, and therefore h_3 is a uniform homotopy. If $x \in R'$, $f_1(x) \in S$ and hence $h_3(x,1) = \phi h_2(x,1) = \phi f_1(x) = f_1(x)$. If $x \in R$, $F_0(x) \in S$ and hence $h_3(x,0) = \phi h_2(x,0) = \phi h_1(x,0) = \phi F_0(x) = F_0(x)$. Hence, if, for $x \in R$, $F_1(x) = h_3(x,1)$, F_1 is an extension of f_1 and is uniformly homotopic to F_0. This completes the proof.

3. Dimension. In our presentation of Hopf's theorems we use the cohomology theory based on some family of coverings, usually the family of the finite coverings, the family of the locally finite coverings, or some family of coverings cofinal with one of these. Since the covering dimension also appears in the statements of these theorems it would be convenient in each case to base both the dimension theory and the cohomology theory on the same family of coverings. On the other hand it is not desirable to assign several dimension numbers to one space. In this section we show that the dimension of a normal space is the same whether based on finite, star-finite, or locally finite coverings of the space.[11]

Accordingly, we temporarily assign three dimension numbers to each space. Thus $\dim_F R \leq n$ means that every finite covering of R has a finite refinement of order $\leq n+1$, $\dim_S R \leq n$ means that every star-finite covering of R has a star-finite refinement of order $\leq n+1$, and $\dim_L R \leq n$ means that every locally finite covering of R has a locally finite refinement of order $\leq n+1$. In Theorem 3.5 we show that if R is a normal space these three definitions of dimension are equivalent. The following four lemmas are used in the proof of this theorem.

LEMMA 3.1. *Let f be a mapping of a closed set R' of a normal space R in the n-sphere S^n. Let $\bar{\sigma}^n$ be a closed n-simplex of a subdivision of S^n and let S^{n-1} be its boundary. If the partial*[12] *mapping $f \mid f^{-1}S^{n-1}$ can be extended to a mapping g of $f^{-1}\bar{\sigma}^n$ into S^{n-1}, then f can be extended to a mapping of R into S^n.*

---

[11] An example considered by J. Dieudonné ([11], p. 67) shows that a normal space can have zero dimension in terms of finite or locally finite coverings but infinite dimension in terms of arbitrary coverings.

[12] If f is a mapping defined on a set A and if B is a subset of A we denote by $f \mid B$ the mapping f restricted to B.

Proof. Since any two mappings of a space into a closed n-cell are uniformly homotopic the mapping g of $f^{-1}\overline{\sigma}^n$ into $\overline{\sigma}^n$ is uniformly homotopic to the partial mapping $f \mid f^{-1}\overline{\sigma}^n$. Therefore, if f_1 is the mapping of R' in $S^n - \sigma^n$ which coincides with g on $f^{-1}\overline{\sigma}^n$ and coincides with f on $f^{-1}(S^n - \sigma^n)$, then f_1 and f are uniformly homotopic mappings of R' into S^n. But, by Tietze's Lemma, the mapping f_1 of R' into the n-cell $S^n - \sigma^n$ can be extended to a mapping of R in $S^n - \sigma^n \subset S^n$. Therefore, by Theorem 2. 1, f can be extended to a mapping of R in S^n.

LEMMA 3. 2. *If N is an arbitrary complex with the natural metric, the covering of N by the stars of its vertices has a refinement $\mathfrak{W}$ such that, in the nerve of the covering $\mathfrak{W}$, the star of each vertex is finite dimensional.*

Proof. Let $\{u_i\}$ be the vertices of N and let $St(u_i)$ be the star of u_i. Let B_i be the boundary of the star of u_i; $B_i = \overline{St(u_i)} - St(u_i)$. Let N^n be the n-dimensional part of N, i. e., the complex consisting of the simplexes of dimension $\leq n$.

We define a sequence of neighborhoods of the simplexes of N as follows: If $\sigma^n = u_i u_j \cdots u_k$ is an n-dimensional simplex of N, let $G_r(\sigma^n)$ be the set of points of N for which $x_i + x_j + \cdots + x_k > 1 - (r+1)(r+2)^{-1} \cdot 2^{-n-2}$, where x_i, x_j, etc., are barycentric coordinates. Let $\bar{G}_r(\sigma^n)$ be the set of points for which $x_i + x_j + \cdots + x_k \geq 1 - (r+1)(r+2)^{-1} \cdot 2^{-n-2}$. Let $G_\omega(\sigma^n)$ be the set of points for which $x_i + x_j + \cdots + x_k > 1 - 2^{-n-2}$, and let $\bar{G}_\omega(\sigma^n)$ be the set of points for which $x_i + x_j + \cdots + x_k \geq 1 - 2^{-n-2}$. The barycentric coordinates are continuous in N. Hence $G_r(\sigma^n)$ and $G_\omega(\sigma^n)$ are open while $\bar{G}_r(\sigma^n)$ and $\bar{G}_\omega(\sigma^n)$ are closed. If K is a closed subcomplex of N we define $G_r(K)$ to be the sum of the neighborhoods G_r of the simplexes of K, and similarly we define $\bar{G}_r(K)$, $G_\omega(K)$ and $\bar{G}_\omega(K)$.

The set $\bar{G}_r(K)$ is a closed set. For let p be an arbitrary point of $N - \bar{G}_r(K)$. It will be shown that p is contained in an open set U which does not intersect $\bar{G}_r(K)$. The point p is in some simplex of $N - K$, say $\sigma^n = u_h \cdots u_i \cdots u_j$. The closure $\overline{\sigma}^n$ of σ^n intersects K in a finite or empty complex $\overline{\sigma}^n \cdot K$. Let $U = G_r(\sigma^n) - \bar{G}_r(\overline{\sigma}^n \cdot K)$. The set $\bar{G}_r(\overline{\sigma}^n \cdot K)$, being a finite sum of closed sets, is closed and hence U is open. The point p is in σ^n and hence in $G_r(\sigma^n)$ but since p is not in $\bar{G}_r(K)$ it is not in $\bar{G}_r(\overline{\sigma}^n \cdot K)$. Hence p is contained in the open set U. Now suppose if possible that there is a point $p_1 \in U \cdot \bar{G}_r(K)$. Then $p_1 \in \bar{G}_r(\sigma^m)$ where σ^m is in K but not in $\overline{\sigma}^n \cdot K$. Let $\sigma^m = u_i \cdots u_j \cdots u_k$ and let $\overline{\sigma}^q$ be the intersection of $\overline{\sigma}^n$ and $\overline{\sigma}^m$, where $\sigma^q = u_i \cdots u_j$. Then $\sigma^q \subset \overline{\sigma}^n \cdot K$ and $q < n$ and $q < m$. Of

course σ^q may be empty. Thus p_1 is in $G_r(\sigma^n)$ and in $\bar{G}_r(\sigma^m)$ but since $\sigma^q \subset \bar{\sigma}^n K$, p_1 is not in $\bar{G}_r(\sigma^q)$. Hence the barycentric coordinates of p_1 satisfy the inequalities

$$x_h + \cdots + x_i + \cdots + x_j > 1 - \frac{r+1}{r+2} \cdot 2^{-n-2} \geq 1 - \frac{r+1}{r+2} \cdot 2^{-q-3},$$

$$x_i + \cdots + x_j + \cdots + x_k \geq 1 - \frac{r+1}{r+2} \cdot 2^{-m-2} \geq 1 - \frac{r+1}{r+2} \cdot 2^{-q-3},$$

$$x_i + \cdots + x_j < 1 - \frac{r+1}{r+2} \cdot 2^{-q-2}.$$

Adding the first two and substracting the third gives

$$x_h + \cdots + x_i + \cdots + x_j + \cdots + x_k > 1,$$

which is impossible. Hence $U \cdot \bar{G}_r(K) = 0$ and hence $\bar{G}_r(K)$ is closed. Similarly $\bar{G}_\omega(K)$ is closed. The sets $G_r(K)$ and $G_\omega(K)$ are sums of open sets and hence are open.

Let $V_i = St(u_i) - \bar{G}_\omega(B_i)$. Thus V_i is an open set. We show that $\mathfrak{V} = \{V_i\}$ is a covering of N. For let p be any point of N. Then p is contained in some simplex $u_i \cdots u_k$, and therefore the barycentric coordinates of p satisfy the equality $x_i + \cdots + x_k = 1$. Let x_i be a largest coordinate of p, i. e., $x_i \geq x_j$ for each u_j of the simplex. Suppose, if possible, that p is not in V_i. Then p is in $\bar{G}_\omega(u_h \cdots u_j)$ where $u_h \cdots u_j$ is some simplex not containing u_i. Let its dimension be n. Then $x_h + \cdots + x_j \geq 1 - 2^{-n-2}$. Since x_i is as large as any of the $n+1$ numbers $x_h, \cdots, x_j$, therefore

$$x_i \geq \frac{1}{n+1}\left(1 - \frac{1}{2^{n+2}}\right) = \frac{2^{n+2}-1}{n+1} \cdot \frac{1}{2^{n+2}} > \frac{1}{2^{n+2}}.$$

Therefore $x_i + x_h + \cdots + x_j > 1$ which is impossible. Therefore p is in V_i. Therefore $\mathfrak{V}$ is a covering of N.

Also $\mathfrak{V}$ is a locally finite covering of N. For let p be any point of N. Let σ^n be the simplex of N which contains p. Then $G_\omega(\sigma^n)$ is an open set containing p. But $G_\omega(\sigma^n) \subset \bar{G}_\omega(\sigma^n)$ which does not intersect V_i unless u_i is a vertex of σ^n. Hence $G_\omega(\sigma^n)$ intersects only a finite number of sets of $\mathfrak{V}$. Hence $\mathfrak{V}$ is locally finite.

Let $H_0 = G_0(N^0)$, $H_1 = G_1(N^1)$ and, if $n > 1$, $H_n = G_n(N^n) - \bar{G}_{n-2}(N^{n-2})$. We show that $\mathfrak{H} = \{H_n\}$ is a locally finite covering of N of order ≤ 2. Let p be any point of N. Then p is in some simplex σ^m of N and hence in some set $\bar{G}_m(N^m)$. Let the smallest m for which

$p \in \bar{G}_m(N^m)$ be $m = n-1$. Then $p \in \bar{G}_{n-1}(\sigma^q)$ for some q-simplex σ^q with $q \leq n-1$. Hence $p \in G_n(\sigma^q) \subset G_n(N^n)$. But p is not in $G_{n-2}(N^{n-2})$. Therefore $p \in H_n$. Therefore $\mathfrak{H}$ is a covering of N. Also p is contained in the open set $G_n(N^n)$ which does not intersect H_m if $m \geq n+2$. Hence $\mathfrak{H}$ is a locally finite covering. Also, since $p \in \bar{G}_{n-1}(N^{n-1})$, p is not in H_m for $m > n$ and, since p is not in $\bar{G}_{n-2}(N^{n-2})$, p is not in H_m for $m \leq n-2$. Hence p is in at most 2 sets of $\mathfrak{H}$ and therefore $\mathfrak{H}$ is of order ≤ 2.

Let $W_{in} = V_i \cdot H_n$ and let $\mathfrak{W} = \{W_{in}\}$. Since $\mathfrak{V}$ and $\mathfrak{H}$ are locally finite coverings of N, $\mathfrak{W}$ is a locally finite covering of N. Since $W_{in} \subset V_i \subset St(u_i)$, $\mathfrak{W}$ is a refinement of the covering of N by the stars of its vertices. We show that the star of each vertex of $N(\mathfrak{W})$ the nerve of $\mathfrak{W}$, is finite dimensional. Let w_{in} be the vertex of $N(\mathfrak{W})$ corresponding to W_{in}. Let p be an arbitrary point of $W_{in} = V_i \cdot H_n$. Then $p \in V_i \cdot G_n(N^n)$ and therefore $p \in V_i \cdot G_n(\sigma^q)$ for some q-simplex with $q \leq n$. For each vertex u_k which is not a vertex of σ^q, $V_k \cdot G_n(\sigma^q) \subset V_k \cdot \bar{G}_\omega(\sigma^q) = 0$. Hence p is not in V_k. Hence p is in at most $q+1$ of the sets of $\mathfrak{V}$. Therefore, since p is in at most two of the sets of $\mathfrak{H}$, p is in at most $2q+2 \leq 2n+2$ sets of $\mathfrak{W}$. Hence the star of w_{in} in $N(\mathfrak{W})$ has dimension at most $2n+1$. This completes the proof of the lemma.

LEMMA 3.3. *If $\mathfrak{U}$ is any locally finite covering of a normal space R there is a locally finite refinement $\mathfrak{U}_2$ of $\mathfrak{U}$ such that the star of each vertex of the nerve of $\mathfrak{U}_2$ is finite dimensional.*

Proof. Let $\mathfrak{U} = \{U_i\}$ and let N be the nerve of $\mathfrak{U}$. By Theorem 1.1 there is a mapping ϕ of R in N such that, for each $U_i \in \mathfrak{U}$, $\phi^{-1}St(u_i) \subset U_i$. Let $\mathfrak{W} = \{W_{in}\}$ be a covering of N defined as in Lemma 3.2. Let $U_{in} = \phi^{-1}(W_{in})$ and let $\mathfrak{U}_2 = \{U_{in}\}$. Since $\mathfrak{W}$ is locally finite so is $\mathfrak{U}_2$. Since $U_{in} = \phi^{-1}W_{in} \subset \phi^{-1}St(u_i) \subset U_i$, $\mathfrak{U}_2$ is a refinement of $\mathfrak{U}$. Let $p \in U_{in}$. Then $\phi(p) \in W_{in}$, and hence $\phi(p)$ is in at most $2n+2$ sets of $\mathfrak{W}$. Hence p is in at most $2n+2$ sets of $\mathfrak{U}_2$. Hence, in the nerve of $\mathfrak{U}_2$, the star of the vertex corresponding to U_{in} has dimension at most $2n+1$. This completes the proof.

Let K be a complex, $\bar{\sigma}^n$ a closed simplex of K and S^{n-1} the boundary of $\bar{\sigma}^n$. Then a mapping ϕ of a space R into K is called *essential* in $\bar{\sigma}^n$ if the partial mapping $\phi \mid \phi^{-1}S^{n-1}$ cannot be extended to a mapping of $\phi^{-1}\bar{\sigma}^n$ into S^{n-1}. If the partial mapping can be extended ϕ is called inessential in $\bar{\sigma}^n$.

A locally finite covering $\mathfrak{U}$ of a normal space R is called a *normal* covering if there is a mapping ϕ of R onto the nerve N of $\mathfrak{U}$ such that each open set

of $\mathfrak{U}$ is the inverse image of the star of the corresponding vertex of N and such that ϕ is essential in every closed simplex of N.

LEMMA 3.4. *Each locally finite covering $\mathfrak{U}$ of a normal space R has a normal refinement which is star-finite if $\mathfrak{U}$ is star-finite and is finite if $\mathfrak{U}$ is finite.*

Proof. Let $\mathfrak{U}$ be any locally finite covering of R and let $\mathfrak{U}_1$ be a refinement of $\mathfrak{U}$ such that the star of each vertex of the nerve N_1 of $\mathfrak{U}_1$, is finite dimensional. If $\mathfrak{U}$ is star-finite or finite, $\mathfrak{U}_1 = \mathfrak{U}$. Let ϕ_1 be a canonical mapping of R into N_1.

We now define, by recursion, a sequence of subcomplexes $\{N_i\}$ of N_1 and a sequence of mappings $\{\phi_i\}$, where ϕ_i is a mapping of R into N_i. Suppose that, starting from N_1 and ϕ_1, we have obtained a subcomplex N_i and a mapping ϕ_i of R into N_i which maps a neighborhood of each point of R continuously into a finite subcomplex of N_i. We call a simplex σ^n of N_i superfluous with respect to ϕ_i if ϕ_i is inessential in $\bar{\sigma}^n$ and if also $\bar{\sigma}^n$ is not on the boundary of a higher dimensional simplex of N_i. Since no superfluous simplex is on a boundary, if we remove all the superfluous simplexes of N_i we get a closed complex N_{i+1}. For each superfluous simplex $\bar{\sigma}^n$ of N_i let the partial mapping $\phi_i \mid \phi_i^{-1}\bar{\sigma}^n$ be replaced by an extension of $\phi_i \mid \phi_i^{-1}(\bar{\sigma}^n - \sigma^n)$ to a mapping of $\phi^{-1}\bar{\sigma}^n$ into $\bar{\sigma}^n - \sigma^n$. Let the resulting mapping of R into N_{i+1} be called ϕ_{i+1}. Because ϕ_{i+1} maps each of the closed sets $\phi^{-1}\bar{\sigma}^n$ continuously into $\bar{\sigma}^n$ or a subcomplex, therefore, if K_i is a finite subcomplex of N_i, ϕ_{i+1} maps $\phi_i^{-1}K_i$ continuously into K_i or a subcomplex. Hence, for each point p of R, a neighborhood mapped by ϕ_i into a finite subcomplex of N_i is mapped continuously by ϕ_{i+1} into a finite subcomplex of N_{i+1}.

Let p be a point of R and W a neighborhood of p mapped by ϕ_1 into a finite subcomplex K of N_1. Then ϕ_i maps W continuously into $K \cdot N_i$. Since K is finite, and since $K \cdot N_{i+1}$ differs from $K \cdot N_i$ at most by omitting a finite number of simplexes of K, it follows that all but a finite number of the complexes $\{K \cdot N_i\}$ are identical. Hence all but a finite number of the mappings mappings $\{\phi_i \mid W\}$ are identical. Hence the sequence $\{\phi_i\}$ has a limit ϕ_ω which is a continuous mapping of R into $N_\omega = \prod_i N_i$.

Let $\bar{\sigma}^n$ be a closed simplex of N_ω and suppose if possible that ϕ_ω is inessential in $\bar{\sigma}^n$. Then clearly each ϕ_i is inessential in $\bar{\sigma}^n$. Hence, by Lemma 3.1, ϕ_i is inessential in each $(n+1)$-simplex of N_i having $\bar{\sigma}^n$ on its boundary, and hence by recursion ϕ_i is inessential in every simplex of

$St_i(\bar{\sigma}^n)$, the star of $\bar{\sigma}^n$ in N_i. The star of each vertex, and hence of each simplex, of N_1 is finite dimensional. Let the dimension of $St_1(\bar{\sigma}^n)$ be m. Hence all m-simplexes of $St_1(\bar{\sigma}^n)$ are superfluous. Hence dim $St_2(\bar{\sigma}^n) = m - 1$ and dim $St_{m-n+1}(\bar{\sigma}^n) = n$. Hence $\bar{\sigma}^n$ is not in N_{m-n+2} and hence not in N_ω. This contradicts the assumption that $\bar{\sigma}^n$ is a closed simplex of N_ω. Hence ϕ_ω is essential in every simplex of N_ω.

It follows in particular that ϕ_ω is a mapping of R *onto* N_ω. We define $\mathfrak{U}_\omega$ to be the covering of R by the inverse images of the stars of the vertices of N_ω. Then, if we let each vertex u_i of N_ω correspond to the open set $U_{\omega i} = \phi_\omega^{-1}St_\omega(u_i)$ of $\mathfrak{U}_\omega$, N_ω becomes the nerve of $\mathfrak{U}_\omega$.

Thus ϕ_ω is a continuous mapping of R onto the nerve of $\mathfrak{U}_\omega$, each open set of $\mathfrak{U}_\omega$ is the inverse image of the star of the corresponding vertex of N_ω, and ϕ_ω is essential in every closed simplex of N_ω. Therefore $\mathfrak{U}_\omega$ is a normal covering of R.

In the clearing off process, the set of points mapped into the star of a vertex never increases. Hence $U_{\omega i} = \phi_\omega^{-1}St_\omega(u_i) \subset \phi_1^{-1}St_1(u_i)$. Since ϕ_1 is canonical, $\phi_1^{-1}St(u_i) \subset U_{1i}$ where U_{1i} is the open set of $\mathfrak{U}_1$ corresponding to the vertex u_i of N_1. Hence $U_{\omega i} \subset U_{1i}$. Therefore $\mathfrak{U}_\omega$ is a refinement of $\mathfrak{U}_1$ and hence a refinement of $\mathfrak{U}$. If $\mathfrak{U}$ is star-finite, $\mathfrak{U}_1 = \mathfrak{U}$ is star-finite and, since $\mathfrak{U}_\omega$ is obtained by shrinking the open sets of $\mathfrak{U}_1$, $\mathfrak{U}_\omega$ is also star-finite. Similarly if $\mathfrak{U}$ is finite so is $\mathfrak{U}_\omega$. This completes the proof of the lemma.

THEOREM 3.5. *If R is a normal space* $\dim_L R = \dim_S R = \dim_F R$.

Proof. It is sufficient to prove that each of the three inequalities, $\dim_L R \leq n$, $\dim_S R \leq n$, $\dim_F R \leq n$, implies both the others. This will be shown by proving that each of these inequalities holds if and only if, for each closed set R' of R, each mapping f of R' into the n-sphere S^n can be extended to a mapping F of R into S^n.

Let $\dim_L R \leq n$ or $\dim_S R \leq n$ or $\dim_F R \leq n$. Let R' be any closed set of R and let f be a mapping of R' into S^n. We assume that S^n is the boundary of an $(n+1)$-simplex σ^{n+1}. By Tietze's Lemma f can be extended to a mapping g of R into $\bar{\sigma}^{n+1}$. Let $\mathfrak{U}_1$ be the covering of R by the inverse images of the stars of the vertices of $\bar{\sigma}^{n+1}$. The nerve N_1 of $\mathfrak{U}_1$ is identified with a subcomplex of $\bar{\sigma}^{n+1}$ so that g is a canonical mapping of R into N_1. Let $\mathfrak{U}_2$ be a locally finite or star-finite or finite refinement of the finite covering $\mathfrak{U}_1$ such that the order of $\mathfrak{U}_2$ is $\leq n+1$. Let ϕ be a canonical

mapping of R into the nerve N_2 of $\mathfrak{U}_2$ and let π be a projection [13] from N_2 to N_1. Since π is continuous on each finite subcomplex of N_2 $\pi\phi$ is a continuous mapping of R into N_1. But, since N_2 has dimension $\leq n$, no simplex of N_2 is mapped into σ^{n+1} and hence $\pi\phi$ maps R into S^n. Let f_1 be the partial mapping $\pi\phi \mid R'$. If p is a point of R', $\phi(p)$ is in the closure of the simplex of N_2 determined by p, and π maps this closed simplex into the closure of the simplex containing $f(p)$. Hence the line segment $[f(p), f_1(p)]$ is contained in one of the closed simplexes of S^n. Let h be the mapping of $R' \times I$ into S^n which maps (p, t) into the point which divides the segment $[f(p), f_1(p)]$ in the ratio $t : 1 - t$. Then h is a uniform homotopy of f and f_1. Therefore, by Theorem 2.1, since f_1 can be extended to a mapping $\pi\phi$ of R into S^n, f can be extended to a mapping F of R into S^n.

Now, conversely, let R be a normal space such that, for each closed set R' of R, each mapping f of R' into S^n can be extended to a mapping F of R into S^n. It follows that no mapping of R into a complex can be essential in any closed $(n+1)$-simplex of the complex. Let $\mathfrak{U}_1$ be any locally finite [star-finite, finite] covering of R. Let $\mathfrak{U}_2$ be a locally finite [star-finite, finite] normal refinement of U_1 (See Lemma 3.4). Then there is a mapping ϕ of R onto the nerve N_2 of U_2 which is essential in every simplex of N_2. Hence N_2 contains no $(n+1)$-simplex. Hence the order of U_2 is $\leq n+1$. Hence $\dim_L R \leq n$ and $\dim_S R \leq n$ and $\dim_F R \leq n$. This completes the proof.

We now define the *dimension* of a normal space R, $\dim R$, to be the common dimension $\dim_L R = \dim_S R = \dim_F R$.

In the proof of Theorem 3.5 we have proved the following:

COROLLARY [14] 3.6. *If R is a normal space, $\dim R \leq n$ if and only if, for each closed set R' of R, each mapping f of R into the n-sphere S^n can be extended to a mapping F of R into S^n.*

4. Čech cohomology groups. This section contains a summary of that

[13] If a covering $\mathfrak{V}$ is a refinement of a covering $\mathfrak{U}$, let each open set $V \in \mathfrak{V}$ be associated with some open set $U \in \mathfrak{U}$ containing V. This induces a mapping of the vertices of the nerve $N(\mathfrak{V})$ into the vertices of the nerve $N(\mathfrak{U})$. This vertex mapping can be extended to a barycentric mapping of the nerve of $\mathfrak{V}$ in the nerve of $\mathfrak{U}$. Such a mapping of one nerve in another is called a projection ([8], p. 157). Barycentric mappings are defined and proved continuous on each finite or locally finite complex by S. Lefschetz ([24], p. 290). It can be shown that any barycentric mapping of one simplicial complex in another is continuous.

[14] This theorem was proved by P. Alexandroff ([2], p. 170) for compact spaces. The extension to normal spaces was discovered independently by E. Hemmingsen.

part of the Čech cohomology theory which is needed in the rest of the paper. The family of coverings on which the cohomology groups are based is not necessarily a family of finite coverings but may be the family of all locally finite coverings or even the family of all coverings of the space. However, in the theorems of this paper, cohomology groups based on the family of all coverings are used only if this family is cofinal with the family of locally finite coverings, i. e., if the space is paracompact.

Let K be an arbitrary closure finite complex in which the boundary of the boundary of each cell is zero. Every simplicial complex satisfies these conditions. An n-chain C^n of K is an expression $a^\alpha \epsilon_\alpha{}^n$ where each a^α is an integer and each $\epsilon_\alpha{}^n$ is an n-cell of K. Since K is closure finite the coboundary $\delta C^n = a^\alpha \eta_\alpha{}^\beta \epsilon_\beta{}^{n+1}$ exists. An n-chain whose coboundary is zero is called an n-cocycle. The n-cocycles of K form an abelian group and the n-coboundaries form a subgroup of the n-cocycles. The difference group of the n-cocycles mod the n-coboundaries is called the n-cohomology group of K and is designated $H^n(K)$. In the Čech cohomology theory, which we shall now summarize, all complexes are assumed to be simplicial.

Let $K = \{\sigma_\alpha{}^n\}$ and $L = \{\tau_\beta{}^n\}$ be two arbitrary simplicial complexes and let f be a simplicial mapping of K into L. Let $f_\beta{}^{n\alpha} = 1, -1$, or 0 according as $\sigma_\alpha{}^n$ is mapped by f onto $\tau_\beta{}^n$ or $-\tau_\beta{}^n$ or is not mapped onto $\tau_\beta{}^n$. If $C^n = a^\beta \tau_\beta{}^n$ is a chain of L let $f^* C^n = a^\beta f_\beta{}^{n+1,\alpha} \sigma_\alpha{}^n$. One easily verifies that $\eta_\gamma{}^\alpha f_\beta{}^{n\gamma} = \bar{\eta}_\beta{}^\gamma f_\gamma{}^{n+1,\alpha}$ where $\eta_\gamma{}^\alpha$ and $\bar{\eta}_\beta{}^\gamma$ are incidence numbers in K and L respectively. Therefore $\delta f^* C^n = a^\beta f_\beta{}^{n\gamma} \eta_\gamma{}^\alpha \sigma_\alpha{}^{n+1} = a^\beta \bar{\eta}_\beta{}^\gamma f_\gamma{}^{n+1,\alpha} \sigma_\alpha{}^{n+1} = f^* \delta C^n$. Hence f^* induces a homomorphism of $H^n(L)$ into $H^n(K)$.

Let the vertices of K be ordered in an arbitrary fashion $u_0, u_1, \cdots, u_n, \cdots, u_\lambda, \cdots$, and let $\{u'_\lambda\}$ and $\{u''_\lambda\}$ be two similar sequences. For each simplex $\sigma_\alpha{}^n$ of K we define $P\sigma_\alpha{}^n$ as follows: If $\sigma_\alpha{}^n = u_{\lambda_0} \cdots u_{\lambda_n}$, where $\lambda_0 < \cdots < \lambda_n$, then $P\sigma_\alpha{}^n = \Sigma_k (-1)^{k-1} u'_{\lambda_0} \cdots u'_{\lambda_k} u''_{\lambda_k} \cdots u''_{\lambda_n}$. The set of all simplexes appearing in the collection $\{P\sigma_\alpha{}^n\}$ together with all their faces, forms a simplicial complex which we denote by PK. The subcomplexes K' and K'' consisting of those simplexes of PK whose vertices have only single or double primes respectively are clearly isomorphic to K under the vertex correspondences $u_\lambda \leftrightarrow u'_\lambda$ and $u_\lambda \leftrightarrow u''_\lambda$. Thus to each chain C^n of K there correspond chains C'^n of K' and C''^n of K'', and conversely.

LEMMA 4.1. *Let Γ^n be a cocycle of PK and let $D_0{}^n$ and $D_1{}^n$ be the parts of Γ^n on K' and K''. Then the corresponding chains $C_0{}^n$ and $C_1{}^n$ of K are cohomologous in K.*

Proof. If $C^n = a^\alpha \sigma_\alpha{}^n$ is a finite chain and $D^n = b^\alpha \sigma_\alpha{}^n$ is any chain we denote by $KI(C^n; D^n)$ the integer $\Sigma_\alpha a^\alpha b^\alpha$. We have $KI(\partial C^{n+1}; D^n) = \Sigma a^\alpha \eta_\beta{}^\alpha b^\beta = KI(C^{n+1}; \delta D^n)$. Hence since Γ^n is a cocycle $KI(\partial C^{n+1}; \Gamma^n) = 0$. One easily computes that $\partial P\sigma_\alpha{}^n = \sigma''_\alpha{}^n - \sigma'_\alpha{}^n - P\partial\sigma_\alpha{}^n$. Then $0 = KI(\partial P\sigma_\alpha{}^n; \Gamma^n) = KI(\sigma''_\alpha{}^n; \Gamma^n) - KI(\sigma'_\alpha{}^n; \Gamma^n) - \Sigma_\beta \eta_\beta{}^\alpha KI(P\sigma_\beta{}^{n-1}; \Gamma^n)$. Let $D_1{}^n = c_1{}^\alpha \sigma''_\alpha{}^n$ and $D_0{}^n = c_0{}^\alpha \sigma'_\alpha{}^n$. Then $KI(\sigma''_\alpha{}^n; \Gamma^n) = c_1{}^\alpha$ and $KI(\sigma'_\alpha{}^n; \Gamma^n) = c_0{}^\alpha$. Let $KI(P\sigma_\beta{}^{n-1}; \Gamma^n) = c^\beta$. Hence, for each α, $0 = c_1{}^\alpha - c_0{}^\alpha - \eta_\beta{}^\alpha c^\beta$. Let C^{n-1} be the chain $c^\beta \sigma_\beta{}^{n-1}$ in K. Then $\delta C^{n-1} = c^\beta \eta_\beta{}^\alpha \sigma_\alpha{}^n = c_1{}^\alpha \sigma_\alpha{}^n - c_0{}^\alpha \sigma_\alpha{}^n = C_1 - C_0$. Hence [14a] $C_1 \sim C_0$ in K, which proves the lemma.

Let R be a space. Let $\mathfrak{U}_\alpha$ and $\mathfrak{U}_\beta$ be coverings of R and let N_α and N_β be their nerves. If $\mathfrak{U}_\beta$ is a refinement of $\mathfrak{U}_\alpha$ there are one or more projections [15] of N_β into N_α. Let $\pi_\alpha{}^\beta$ and $\pi'_\alpha{}^\beta$ be two such projections. It can be shown, as for example by Čech ([8], p. 159), that there is a simplicial mapping f of PN_β into N_α such that $f \mid N'_\beta$ and $f \mid N''_\beta$ correspond to $\pi_\alpha{}^\beta$ and $\pi'_\alpha{}^\beta$. Let $\Gamma_\alpha{}^n$ be a cocycle of N_α. Then $f^*\Gamma_\alpha{}^n$ is a cocycle of $N_\beta \times I$. It follows from Lemma 4.1 that $\pi^*{}_\alpha{}^\beta \Gamma_\alpha{}^n \sim \pi'^*{}_\alpha{}^\beta \Gamma_\alpha{}^n$ in N_β.

A collection $\Gamma^n = \{\Gamma_\alpha{}^n\}$ of one or more cocycles, one on each of the nerves of some of the coverings of R, is called a cocycle of R if the following condition is satisfied: If $\Gamma_\alpha{}^n$ and $\Gamma_\beta{}^n \in \Gamma^n$ there is a common refinement $\mathfrak{U}_\gamma$ of $\mathfrak{U}_\alpha$ and $\mathfrak{U}_\beta$, such that $\pi^*{}_\alpha{}^\gamma \Gamma_\alpha{}^n \sim \pi^*{}_\beta{}^\gamma \Gamma_\beta{}^n$ in N_γ. As shown above, the choice of the projection $\pi^*{}_\alpha{}^\gamma$ and $\pi^*{}_\beta{}^\gamma$ is irrelevant.

Two cocycles $\Gamma_1{}^n$ and $\Gamma_2{}^n$ of R are called cohomologous, $\Gamma_1{}^n \sim \Gamma_2{}^n$, if for each $\Gamma_{1\alpha}{}^n \in \Gamma_1{}^n$ and $\Gamma_{2\beta}{}^n \in \Gamma_2{}^n$ there is a common refinement $\mathfrak{U}_\gamma$ of $\mathfrak{U}_\alpha$ and $\mathfrak{U}_\beta$ such that $\pi^*{}_\alpha{}^\gamma \Gamma_{1\alpha}{}^n \sim \pi^*{}_\beta{}^\gamma \Gamma_{2\beta}{}^n$ in N_γ. The sum $\Gamma_1{}^n + \Gamma_2{}^n$ of the cocycles of R is defined to be any cocycle Γ^n of R such that if $\Gamma_{1\alpha}{}^n \in \Gamma_1{}^n$, $\Gamma_{2\beta}{}^n \in \Gamma_2{}^n$, and $\Gamma_\gamma{}^n \in \Gamma^n$, then there is a common refinement $\mathfrak{U}_\delta$ of $\mathfrak{U}_\alpha$, $\mathfrak{U}_\beta$, and $\mathfrak{U}_\gamma$ such that $\pi^*{}_\gamma{}^\delta \Gamma_\gamma{}^n \sim \pi^*{}_\alpha{}^\delta \Gamma_{1\alpha}{}^n + \pi^*{}_\beta{}^\delta \Gamma_{2\beta}{}^n$ in N_δ. The sum $\Gamma_1{}^n + \Gamma_2{}^n$ exists and is unique up to cohomology.

A cocycle Γ^n of R is determined up to cohomology by any $\Gamma_\alpha{}^n \in \Gamma^n$. Also $\Gamma^n \sim 0$ if for some $\Gamma_\alpha{}^n \in \Gamma^n$ there is a refinement $\mathfrak{U}_\beta$ of $\mathfrak{U}_\alpha$ such that $\pi^*{}_\alpha{}^\beta \Gamma_\alpha{}^n \sim 0$ in N_β.

The classes of cohomologous n-cocycles of R form a group $H^n(R)$, the n-dimensional cohomology group of R.

Let $Z = \{\mathfrak{U}_\alpha\}$ be a family of coverings of R subject to the condition that

---

[14a] For the convenience of the printer we depart from the Whitney [28] notation by using a tilde instead of a reversed tilde for "is cohomologous to." Since only cohomology is used in this paper, no misunderstanding need arise.

[15] See note 13.

any two coverings $\mathfrak{U}_\alpha$ and $\mathfrak{U}_\beta \in Z$ have a common refinement $\mathfrak{U}_\gamma \in Z$. A cocycle Γ^n of R is said to be in Z if for some $\Gamma_\alpha{}^n \in \Gamma^n$ the corresponding covering $\mathfrak{U}_\alpha \in Z$. A cocycle Γ^n in Z is called ~ 0 in Z if for each $\Gamma_\alpha{}^n \in \Gamma^n$ for which the corresponding covering $\mathfrak{U}_\alpha \in Z$ there is a refinement $\mathfrak{U}_\beta \in Z$ of $\mathfrak{U}_\alpha$ such that $\pi^*{}_\alpha{}^\beta \Gamma_\alpha{}^n \sim 0$ in N_β. Let $Z \subset Z'$. Then if Γ^n is a cocycle in Z it is also in Z', and, if $\Gamma^n \sim 0$ in Z, $\Gamma^n \sim 0$ in Z'.

The classes of n-cycles cohomologous in Z form a group $H_Z{}^n(R)$, the n-dimensional cohomology group of the family Z of coverings of R. In particular there are the cohomology groups $H_F{}^n(R)$, $H_S{}^n(R)$, and $H_L{}^n(R)$ of the families of finite, star-finite, and locally finite coverings respectively. The cohomology group $H_F{}^n(R)$ of the finite coverings of R is the usual Čech n-dimensional cohomology group of R.

If Z is a cofinal subfamily of Z', i. e., if every covering $\mathfrak{U} \in Z'$ has a refinement $\mathfrak{V} \in Z$, then $H_Z{}^n(R) = H_{Z'}{}^n(R)$. If Z is a cofinal family of coverings of R, i. e., if every covering $\mathfrak{U}$ of R has a refinement $\mathfrak{V} \in Z$ then $H_Z{}^n(R) = H^n(R)$.

Let R and S be two topological spaces and let f be a continuous mapping of R into S. If $\mathfrak{V}_\alpha$ is any covering of S some of the open sets of $\mathfrak{V}_\alpha$ are mapped by f^{-1} into open sets (not necessarily all distinct) which form a covering $\mathfrak{U}_\alpha$ of R. Let ϕ be the barycentric of $N(\mathfrak{U}_\alpha)$ into $N(\mathfrak{V}_\alpha)$ which sends the vertex of $N(\mathfrak{U}_\alpha)$ corresponding to the open set $f^{-1}(V) \in \mathfrak{U}_\alpha$ into the vertex of $N(\mathfrak{V}_\alpha)$ corresponding to $V \in \mathfrak{V}_\alpha$. If $\Gamma^n = \{\Gamma_\alpha{}^n\}$ is any cocycle of S let $f^*\Gamma^n = \{\phi^*\Gamma_\alpha{}^n\}$. We show that $f^*\Gamma^n$ is a cocycle of R. Let $\mathfrak{V}_\gamma$ be any refinement of $\mathfrak{V}_\alpha$. If $\pi_\alpha{}^\gamma$ is a projection from $N(\mathfrak{V}_\gamma)$ to $N(\mathfrak{V}_\alpha)$ then $\pi'_\alpha{}^\gamma = \phi^{-1}\pi_\alpha{}^\gamma\phi$ is a projection from $N(\mathfrak{U}_\gamma)$ to $N(\mathfrak{U}_\alpha)$ such that $\phi\pi'_\alpha{}^\gamma$ is the same mapping of $N(\mathfrak{U}_\gamma)$ into $N(\mathfrak{V}_\alpha)$ as $\pi_\alpha{}^\gamma\phi$. Therefore $\pi^{*\prime}{}_\alpha{}^\gamma\phi^*\Gamma_\alpha{}^n = \phi^*\pi^*{}_\alpha{}^\gamma\Gamma_\alpha{}^n$. The inverse cocycle mapping $\pi^{*\prime}{}_\alpha{}^\gamma$ is determined up to cohomology independently of the choice of the projection. Hence, for each $\pi_\alpha{}^\gamma$, $\pi^*{}_\alpha{}^\gamma\phi^*\Gamma_\alpha{}^n \sim \phi^*\pi^*{}_\alpha{}^\gamma\Gamma_\alpha{}^n$. Now if $\phi^*\Gamma_\alpha{}^n$ and $\phi^*\Gamma_\beta{}^n \in f^*\Gamma^n$, then $\Gamma_\alpha{}^n$ and $\Gamma_\beta{}^n \in \Gamma^n$ and hence there is a common refinement $\mathfrak{V}_\gamma$ of $\mathfrak{V}_\alpha$ and $\mathfrak{V}_\beta$ such that $\pi^*{}_\alpha{}^\gamma\Gamma_\alpha{}^n \sim \pi^*{}_\beta{}^\gamma\Gamma_\beta{}^n$. Therefore $\pi^*{}_\alpha{}^\gamma\phi^*\Gamma_\alpha{}^n \sim \phi^*\pi^*{}_\alpha{}^\gamma\Gamma_\alpha{}^n \sim \phi^*\pi^*{}_\beta{}^\gamma\Gamma_\beta{}^n \sim \pi^*{}_\beta{}^\gamma\phi^*\Gamma_\beta{}^n$, which shows that $f^*\Gamma^n$ is a cocycle. Similarly using the fact that ϕ^* and π^* commute up to cohomology one can show that if $\Gamma^n \sim 0$ then $f^*\Gamma^n \sim 0$ and if $\Gamma^n = \Gamma'^n \pm \Gamma''^n$ then $f^*\Gamma^n = f^*\Gamma^n = f^*\Gamma'^n \pm f^*\Gamma''^n$. Thus f induces a homomorphism of $H^n(S)$ into $H^n(R)$.

Let R' be a closed subspace of the space R. If $\mathfrak{U}_\alpha$ is a covering of R the intersections with R' of the open sets of $\mathfrak{U}_\alpha$ form a covering $\mathfrak{U}'_\alpha$ of R' and every covering $\mathfrak{U}'_\alpha$ of R' can be so obtained from a covering $\mathfrak{U}_\alpha$ of R. The nerve $N'_\alpha = N(\mathfrak{U}'_\alpha)$ can be considered as a closed subcomplex of $N'_\alpha = N(U'_\alpha)$.

Let $\Gamma^n = \{\Gamma_\alpha{}^n\}$ be a cocycle of R. If each $\Gamma_\alpha{}^n$ is in $N_\alpha - N'_\alpha$ and if the cohomologies $\pi^*{}_\alpha{}^\gamma\Gamma_\alpha{}^n \sim \pi^*{}_\beta{}^\gamma\Gamma_\beta{}^n$ can be satisfied in $N_\gamma - N'_\gamma$ we say that Γ^n is in $R - R'$. For such a Γ^n we say that $\Gamma^n \sim 0$ in $R - R'$ if the cohomologies $\pi^*{}_\alpha{}^\beta\Gamma_\alpha{}^n \sim 0$ can be satisfied in $N_\beta - N'_\beta$.

Let $\Gamma'^n = \{\Gamma'_\alpha{}^n\}$ be a cocycle of the space R'. Then $\Gamma'_\alpha{}^n$ is a chain of N_α such that $\delta\Gamma'_\alpha{}^n$ is in $N_\alpha - N'_\alpha$. The cocycle $\delta\Gamma'_\alpha{}^n$ is determined up to cohomology in $N_\alpha - N'_\alpha$ by the cohomology class of $\Gamma'_\alpha{}^n$. For if $\Gamma'_{1\alpha}{}^n \sim \Gamma'_{2\alpha}{}^n$ in N'_α there is a chain C^{n-1} in N'_α and a chain C^n in $N_\alpha - N'_\alpha$ such that $\delta C^{n-1} = \Gamma'_{1\alpha}{}^n - \Gamma'_{2\alpha}{}^n - C^n$ in N_α. Hence $0 = \delta\delta C^{n-1} = \delta\Gamma'_{1\alpha}{}^n - \delta\Gamma'_{2\alpha}{}^n - \delta C^n$. Hence $\delta\Gamma'_{1\alpha}{}^n \sim \delta\Gamma'_{2\alpha}{}^n$ in $N_\alpha - N'_\alpha$. We now show that $\Gamma^{n+1} = \{\delta\Gamma'_\alpha{}^n\}$ is a cocycle of R in $R - R'$. Let $\pi'_\alpha{}^\gamma$ be the projection from N'_γ to N'_α, etc. If $\delta\Gamma'_\alpha{}^n$ and $\delta\Gamma'_\beta{}^n$ are in Γ^{n+1} then $\Gamma'_\alpha{}^n$ and $\Gamma'_\beta{}^n$ are in Γ'^n. Hence there is a refinement $\mathfrak{U}'_\gamma$ of $\mathfrak{U}'_\alpha$ and $\mathfrak{U}'_\beta$ such that $\pi'^*{}_\alpha{}^\gamma\Gamma'_\alpha{}^n \sim \pi'^*{}_\beta{}^\gamma\Gamma'_\beta{}^n$ in N'_γ. It can be assumed that $\mathfrak{U}'_\gamma$ is the intersection with R' of $\mathfrak{U}_\gamma$, a refinement of $\mathfrak{U}_\alpha$ and $\mathfrak{U}_\beta$. Then $\delta\pi'^*{}_\alpha{}^\gamma\Gamma'_\alpha{}^n \sim \delta\pi'^*{}_\beta{}^\gamma\Gamma'_\beta{}^n$ in $N_\gamma - N'_\gamma$. Now $\pi^*{}_\alpha{}^\gamma\Gamma'_\alpha{}^n = \pi'^*{}_\alpha{}^\gamma\Gamma'_\alpha{}^n + C^n$ where C^n is a chain of $N_\gamma - N'_\gamma$ and hence $\delta\pi^*{}_\alpha{}^\gamma\Gamma'_\alpha{}^n \sim \delta\pi'^*{}_\alpha{}^\gamma\Gamma'_\alpha{}^n$ in $N_\gamma - N'_\gamma$. Therefore, since $\delta\pi^* = \pi^*\delta$, $\pi^*{}_\alpha{}^\gamma\delta\Gamma'_\alpha{}^n \sim \pi^*{}_\beta{}^\gamma\delta\Gamma'_\beta{}^n$ in $N_\gamma - N'_\gamma$. Therefore Γ^{n+1} is a cocycle of R in $R - R'$. We denote Γ^{n+1} by $\Psi(\Gamma'^n)$.

Let R' be a closed set of R and let S' be a closed set of S. If F is a continuous mapping of R into S which maps R' into S' then F^* can be seen to map cocycles of $S - S'$ into cocycles of $R - R'$ and to map cocycles which are cohomologous in $S - S'$ into cocycles which are cohomologous in $R - R'$. Thus F^* induces a homomorphism of the n-cohomology group of $S - S'$ into the n-cohomology group of $R - R'$. Let Δ'^n be any n-cocycle of S' and let $\mathfrak{V}_\alpha$ be a covering of S such that $\Delta'_\alpha{}^n$ exists in the nerve of the corresponding covering $\mathfrak{V}'_\alpha$ of S'. The inverse images of the open sets of $\mathfrak{V}_\alpha$ and $\mathfrak{V}'_\alpha$ form coverings of R and R' with nerves N and N' such that $N'_\alpha \subset N_\alpha$. There are corresponding mappings Φ and $\phi = \Phi \mid N'_\alpha$ of N_α and N'_α into the nerves $N(\mathfrak{V}_\alpha)$ and $N(\mathfrak{V}'_\alpha) \subset N(\mathfrak{V}_\alpha)$ respectively. Then $\Phi^*\Delta'_\alpha{}^n = \phi^*\Delta'_\alpha{}^n + C^n$, where C^n is a chain of $N - N'$. Hence $\delta\Phi^*\Delta'_\alpha{}^n \sim \delta\phi^*\Delta'_\alpha{}^n$ in $N - N'$. But $\delta\Phi^*\Delta'_\alpha{}^n = \Phi^*\delta\Delta'_\alpha{}^n$. Hence $\delta\phi^*\Delta'_\alpha{}^n \sim \Phi^*\delta\Delta'_\alpha{}^n$. Hence $\Psi f^*\Delta'^n = \{\delta\phi^*\Delta'_\alpha{}^n\} \sim \{\Phi^*\delta\Delta'_\alpha{}^n\} = F^*\Psi\Delta'^n$ in $R - R'$. Thus, finally, $\Psi f^*\Delta'^n \sim F^*\Psi\Delta'^n$ in $R - R'$.

5. Hopf's extension theorem. The n-dimensional cohomology group of the n-dimensional sphere S^n is cyclic infinite. Let Δ^n be a fixed cocycle in the generator of this group. We may assume that Δ^n is a cocycle of S^n in the family of finite coverings of S^n. We call Δ^n the fundamental cocycle of S^n.

LEMMA 5.1. *Let R be a normal space with* $\dim R \leq n + 1$. *Let R' be a*

closed subspace of R and let f be a mapping of R' into S^n. Let Z be a family of coverings of R such that all finite coverings of R are in Z and such that every covering in Z has a locally finite refinement. Then f can be extended to a mapping F of R into S^n if and only if $\Psi f^*\Delta^n \sim 0$ in $R - R'$ in Z.

Proof. 1) Let F be an extension of f. Let $\Delta^n = \{\Delta_a{}^n\}$. Then $F^*\Delta^n = \{\Phi^*\Delta_a{}^n\}$ is a cocycle of R and $f^*\Delta^n = \{\phi^*\Delta_a{}^n\}$ is a cocycle of R'. Let $\Delta_a{}^n \varepsilon \Delta^n$ and let $\mathfrak{V}_a$ be a finite covering of S^n. Then $\mathfrak{U}_a = F^{-1}\mathfrak{V}_a$ is a finite covering of R. Then $\Phi^*\Delta_a{}^n$ is a cocycle of $N(\mathfrak{U}_a)$ and $\Phi^*\Delta_a{}^n = \phi^*\Delta_a{}^n + C_a{}^n$ where $C_a{}^n$ is a chain in $N(\mathfrak{U}_a) - N(\mathfrak{U}'_a)$. Therefore $\delta\phi^*\Delta_a{}^n = \delta(-C_a{}^n)$. This shows that $\Psi f^*\Delta^n \sim 0$ in $R - R'$ in the family of finite coverings of R. Therefore, since Z contains the finite coverings, $\Psi f^*\Delta^n \sim 0$ in $R - R'$ in Z.

2) Let f be a mapping of R' into S^n such that $\Psi f^*\Delta^n \sim 0$ in $R - R'$ in Z. We may assume that S^n is the boundary of an $(n+1)$-simplex σ^{n+1}. Then S^n is a complex. Let $\mathfrak{V}_0$ be the covering of S^n by the stars of its vertices. We may assume that Δ^n is a cocycle of the generating element of $H^n(S^n)$ such that, for some cocycle $\Delta_0{}^n$ of $N(\mathfrak{V}_0)$, $\Delta_0{}^n \varepsilon \Delta^n$. Let $\mathfrak{U}'_0 = f^{-1}(\mathfrak{V}_0)$ and let $\mathfrak{U}_0$ be an extension of $\mathfrak{U}'_0$ to a covering of R. Then $N'_0 = N(\mathfrak{U}'_0)$ is isomorphic to a subcomplex of S^n. We identify N'_0 with this subcomplex so that $N'_0 \subset S^n$.

Let $\Gamma^n = f^*\Delta^n$. Let $\mathfrak{U}_\beta \varepsilon Z$ be a refinement of $\mathfrak{U}_0$ such that $\delta\pi^*{}_0{}^\beta\Gamma_0{}^n \sim 0$ in $N_\beta - N'_\beta$. Let $\mathfrak{U}_\alpha$ be a locally finite refinement of $\mathfrak{U}_\beta$ of order $\leq n+2$. Then $\delta\pi_0{}^{*\alpha}\Gamma_0{}^n \sim 0$ in $N_\alpha - N'_\alpha$. Therefore $\pi_0{}^{*\alpha}\Gamma_0{}^n$ is part of a cocycle in N_α. The proof of Hopf's extension theorem for complexes ([28], p. 54) shows that the mapping $\pi_0{}^\alpha$ of N'_α into S^n can be extended to a mapping $\bar{\pi}$ of N_α into S^n which is continuous on every finite subcomplex of N_α.

Let ϕ be a canonical mapping of R in N_α. Then $\bar{\pi}\phi$ is a continuous mapping of R into S^n. Let f_0 be the partial mapping $\bar{\pi}\phi \mid R'$. Each point $x \varepsilon R'$ is mapped by ϕ in the closure of the simplex of N'_α determined by x. This closed simplex is mapped by $\bar{\pi} = \pi_0{}^\alpha$ into the closure of the simplex of N'_0 determined by x, i. e., into the closure of the simplex of S^n containing $f(x)$. Therefore $f_0(x) = \bar{\pi}\phi(x)$ is in the same closed simplex as $f(x)$. Let h be the mapping of $R' \times I$ in S^n which maps (x,t) in the point dividing the segment $[f_0(x), f(x)]$ in the ratio $t : 1-t$. Then h is a uniform homotopy. Therefore by Borsuk's theorem (Theorem 2.1), since f_0 is uniformly homotopic to f, and since f_0 can be extended to a mapping $\bar{\pi}\phi$ of R into S^n, f can be extended to a mapping F of R into S^n.

THEOREM 5.2. (*Hopf's extension theorem*) *Let R be a paracompact*

normal space with $\dim R \leq n+1$. *Let* R' *be a closed set of* R *and let* f *be a mapping of* R' *into the* n*-sphere* S^n. *Then* f *can be extended to a mapping* F *of* R *into* S^n *if and only if* $\Psi f^*\Delta^n \sim 0$ *in* $R - R'$.

Proof. Since R is paracompact every covering of R has a locally finite refinement. If in Lemma 5.1 we let Z be the family of all open coverings of R we get Theorem 5.2 as an immediate consequence.

Hopf's extension theorem takes the following form if we use the ordinary Čech cocycles based on finite coverings.

THEOREM 5.3. (*Hopf's extension theorem with finite coverings*). *Let* R *be a normal space with* $\dim R \leq n+1$. *Let* R' *be a closed subspace of* R *and let* f *be a mapping of* R' *into the* n*-sphere* S^n. *Then* f *can be extended to a mapping* F *of* R *into* S^n *if and only if* $\Psi f^*\Delta^n \sim 0$ *in* $R - R'$ *in the family of finite coverings of* R.

Proof. This theorem is obtained as an immediate consequence of Lemma 5.1 by letting Z be the family of finite coverings of R.

6. A class of coverings of $R \times I$. A certain class of coverings of the product space $R \times I$ plays an important role in the proof of Hopf's classification theorem. It is a generalization of the product covering which consists of the products of the sets of a covering of R by the sets of a covering of I. Let $\mathfrak{V} = \{V_{ai}\}$ be a covering of $R \times I$ by open sets each of which is a product set $V_{ai} = U_a \times W_{ai}$ of an open set of a locally finite covering $\mathfrak{U} = \{U_a\}$ by an open set (interval) of a covering $\mathfrak{W}_a$, corresponding to U_a, where $\mathfrak{W}_a = \{W_{ai}\}$ is an irreducible, and hence finite, covering of I by at least two intervals. It is assumed that for each irreducible covering, $\mathfrak{W}_a = \{W_{a0}, W_{a1}, W_{a2}, \cdots, W_{ai}, \cdots, W_{a,q(a)}\}$, each two successive intervals intersect, no other pairs intersect, $0 \in W_{a0}$, and $1 \in W_{a,q(a)}$ where $q(a) \neq 0$. Such coverings $\mathfrak{V}$ of $R \times I$ will be called P-coverings.

Let N be the nerve of a P-covering $\mathfrak{V} = \{V_{ai}\}$ and let v_{ai} be the vertex of N corresponding to $V_{ai} = U_a \times W_{ai}$. The intersection of $\mathfrak{V}$ with $R \times 0$ and $R \times 1$ determine a covering $\mathfrak{U}' = \{U_a \times 0\}$ of $R \times 0$ and a covering $\mathfrak{U}'' = \{U_a \times 1\}$ of $R \times 1$ whose nerves N' and N'' we identify with the subcomplexes of N having vertices $\{v_{a0}\}$ and $\{v_{a,q(a)}\}$ respectively. Under the correspondence of the vertices u_a of $N(\mathfrak{U})$ to v_{a0} and $v_{a,q(a)}$, the complex $N(\mathfrak{U})$ corresponds to the isomorphic complexes N' and N'', and the simplexes σ^n, chains C^n, etc. of $N(\mathfrak{U})$ correspond to the simplexes σ'^n, σ''^n, chains C'^n, C''^n, etc. of N' and N'' respectively.

Let the vertices of N be ordered as follows: Let $y_{\alpha i} \in I$ be chosen for each set $W_{\alpha i}$ so that, if $i = 0$, $y_{\alpha i} = 0$ and, if $i > 0$, $y_{\alpha i} \in W_{\alpha,i-1} \cdot W_{\alpha i}$. Thus for each α the points $y_{\alpha i}$ are ordered like the subscripts i. We say that $v_{\alpha i} < v_{\beta j}$ if either $y_{\alpha i} < y_{\beta j}$ or, when $y_{\alpha i} = y_{\beta j}$, $\alpha < \beta$.

Corresponding to each n-chain C^n of $N(\mathfrak{U})$ we define a deformation chain DC^n of N as follows: Let $\sigma^n = u_\alpha u_\beta \cdots u_\gamma$. Then $D\sigma^n$ is defined to be the sum of all $n+1$-simplexes σ^{n+1} of the form $v_{\delta,m-1}v_{\alpha i}v_{\beta j} \cdots v_{\delta m} \cdots v_{\gamma k}$ (where δ may be any of the subscripts $\alpha, \beta, \cdots \gamma$), such that for each vertex $v_{\epsilon r}$, with, $\epsilon \neq \delta$, $v_{\epsilon r} < v_{\delta m} < v_{\epsilon,r+1}$, the last relation being assumed trivially true if $v_{\epsilon,r+1}$ does not exist. Each such simplex exists in N if its vertices exist. For, since $v_{\epsilon r} < v_{\delta m}$, $y_{\epsilon r} \leq y_{\delta m}$ and, if $v_{\epsilon,r+1}$ exists, $y_{\delta m} \leq y_{\epsilon,r+1}$. Thus either $y_{\delta m}$ is between two points $y_{\epsilon r}$ and $y_{\epsilon,r+1}$ of the interval $W_{\epsilon r}$ or it is beyond a point $y_{\epsilon r}$ of a final interval $W_{\epsilon r} = W_{\epsilon,q(\epsilon)}$. In either case $y_{\delta m} \in W_{\epsilon r}$. Let $p \in R$ be a point of the intersection $U_\alpha \cdot U_\beta \cdots U_\gamma$. Then $(p, y_{\delta m}) \in U_\epsilon \times W_{\epsilon r} = V_{\epsilon r}$. Also, since $y_{\delta m} \in W_{\delta,m-1} \cdot W_{\delta m}$, the point $(p, y_{\delta m}) \in V_{\delta,m-1} \cdot V_{\delta m}$. Hence $(p, y_{\delta m}) \in V_{\delta,m-1} \cdot V_{\alpha i} \cdot V_{\beta j} \cdots V_{\delta m} \cdots V_{\gamma k}$. Therefore the simplex $v_{\delta,m-1}v_{\alpha i}v_{\beta j} \cdots v_{\delta m} \cdots v_{\gamma k}$ exists in N. Thus $D\sigma^n$ is an $n+1$-chain of N. If $C^n = a^\alpha \sigma_\alpha{}^n$ is an n-chain of $N(\mathfrak{U})$, DC^n is defined to be the $n+1$-chain $DC^n = a^\alpha D\sigma_\alpha{}^n$ of N.

LEMMA 6.1. *In the nerve of a P-covering of* $R \times I$,

$$\partial D\sigma^n = \sigma''^n - \sigma'^n - D\partial\sigma^n.$$

Proof. Let $\sigma^n = u_\alpha u_\beta \cdots u_\gamma$ and let $\sigma^{n+1} = v_{\delta,m-1}v_{\alpha i}v_{\beta j} \cdots v_{\delta m} \cdots v_{\gamma k}$ be a simplex of $D\sigma^n$. Then $\partial\sigma^{n+1} = v_{\alpha i}v_{\beta j} \cdots v_{\delta m} \cdots v_{\gamma k} - v_{\alpha i}v_{\beta j} \cdots v_{\delta,m-1} \cdots v_{\gamma k} - v_{\delta,m-1}v_{\beta j} \cdots v_{\delta m} \cdots v_{\gamma k} + \cdots$. The first two simplexes of $\delta\sigma^{n+1}$, where all the subscripts $\alpha, \beta, \cdots$ of σ^n are represented, are called simplexes over σ^n, while the others are called simplexes over $\partial\sigma^n$. If $\sigma_1{}^n = v_{\alpha i}v_{\beta j} \cdots v_{\gamma k}$ is any n-simplex over σ^n such that, for each pair of its vertices $v_{\epsilon r}$ and $v_{\delta m}$, $v_{\delta m} < v_{\epsilon,r+1}$ and $v_{\epsilon r} < v_{\delta,m+1}$ then $\sigma_1{}^n$ is on the boundary of exactly two simplexes of $D\sigma^n$ and with opposite signs, or $\sigma_1{}^n = \sigma'^n$ or $\sigma_1{}^n = \sigma''^n$. For, if $\sigma_1{}^n \neq \sigma'^n$, and if $v_{\delta m}$ is the largest vertex of $\sigma_1{}^n$ (in the sense of $>$), then $m \neq 0$ and $v_{\delta,m-1}v_{\alpha i} \cdots v_{\delta m} \cdots v_{\gamma k}$ is the only simplex of $D\sigma^n$ having $\sigma_1{}^n$ on its boundary with coefficient $+1$. And, if $\sigma_1{}^n \neq \sigma''^n$, there is a least, say $v_{\epsilon,r+1}$, of the vertices over vertices of σ^n such that $v_{\epsilon,r+1}$ is greater than all the vertices of $\sigma_1{}^n$, and $v_{\epsilon r}v_{\alpha i} \cdots v_{\epsilon,r+1} \cdots v_{\gamma k} = -v_{\epsilon,r+1}v_{\alpha i} \cdots v_{\epsilon r} \cdots v_{\gamma k}$ is the only simplex of $D\sigma^n$ having $\sigma_1{}^n$ on its boundary with coefficient -1. Hence if $\sigma_1{}^n \neq \sigma'^n$ and $\neq \sigma''^n$, $\sigma_1{}^n$ has coefficient 0 in $\partial D\sigma^n$. And σ'^n is on the

boundary of only one simplex of $D\sigma^n$, where it has coefficient -1, and σ''^n is on the boundary of only one simplex of $D\sigma^n$, where it has coefficient $+1$. Hence $\partial D\sigma^n = \sigma''^n - \sigma'^n +$ simplexes over $\partial\sigma^n$.

If u_ϵ is the ρ-th vertex of σ^n then a typical simplex over $\partial\sigma^n$ of $\partial\sigma^{n+1}$ is $(-1)^\rho v_{\delta,m-1}v_{\alpha i}\cdots\bar{v}_\epsilon\cdots v_{\delta m}\cdots v_{\gamma k}$ where v_ϵ means that the vertex with subscript ϵ is omitted. This is a simplex of the deformation chain of the oriented $n-1$-simplex $(-1)^\rho u_\alpha\cdots\bar{u}_\epsilon\cdots u_\delta\cdots u_\gamma$. The boundary of σ^n is $\partial\sigma^n = \bar{u}_\alpha u_\beta\cdots u_\gamma - u_\alpha\bar{u}_\beta\cdots u_\gamma + \cdots + (-1)^{\rho+1}u_\alpha\cdots\bar{u}_\epsilon\cdots u_\gamma + \cdots$. Thus each simplex of $\partial\sigma^{n+1}$ over $\partial\sigma^n$ is a simplex of $D\partial\sigma^n$ with orientation reversed. Further if $(-1)^{\rho+1}v_{\delta,m-1}v_{\alpha i}\cdots\bar{v}_\epsilon\cdots v_{\delta m}\cdots v_{\gamma k}$ is any simplex of $D\partial\sigma^n$ there is one and only one choice of r, namely the largest for which $a_{\epsilon r} < a_{\delta m}$, such that $v_{\delta,m-1}v_{\alpha i}\cdots v_{\epsilon r}\cdots v_{\delta m}\cdots v_{\gamma k}$ is a simplex of $D\sigma^n$. Therefore $\partial D\sigma^n = \sigma''^n - \sigma'^n - D\partial\sigma^n$, which completes the proof.

LEMMA 6.2. *If $\mathfrak{V}$ is a P-covering over $\mathfrak{U}$ of $R\times I$, and if D_0^n and D_1^n are the parts on N' and N'' of an n-cocycle Γ^n of $N(\mathfrak{V})$ then the corresponding cocycles C_0^n and C_1^n are cohomologous in $N(\mathfrak{U})$.*

Proof. Let $C_1^n = a''^\alpha\sigma_\alpha^n$ and $C_0^n = a'^\alpha\sigma_\alpha^n$. If σ_β^{n-1} is any $n-1$-simplex of $N(\mathfrak{U})$, let $b^\beta = KI(D\sigma_\beta^{n-1};\Gamma^n)$. Now since $\delta\Gamma^n = 0$, $0 = KI(D\sigma_\alpha^n;\delta\Gamma^n) = KI(\partial D\sigma_\alpha^n;\Gamma^n)$. Hence, by Lemma 6.1, $0 = KI(\sigma''_\alpha{}^n;\Gamma^n) - KI(\sigma'_\alpha{}^n;\Gamma^n) - KI(D\partial\sigma_\alpha^n,\Gamma^n) = a''^\alpha - a'^\alpha - \eta_\beta{}^\alpha b^\beta$. Therefore $a''^\alpha\sigma_\alpha^n - a'^\alpha\sigma_\alpha^n - \eta_\beta{}^\alpha b^\beta\sigma_\alpha^n = 0$ in $N(\mathfrak{U})$, i.e., $C_1^n - C_0^n - \delta b^\beta\sigma_\beta^{n-1} = 0$. Hence $C_1^n \sim C_0^n$ in $N(\mathfrak{U})$.

LEMMA 6.3. *Let R be a paracompact normal space and let Γ^n be an n-cocycle of $R\times I$. Then if D_0^n and D_1^n are the parts of Γ^n in $R\times 0$ and $R\times 1$ the corresponding cocycles C_0^n and C_1^n of R are cohomologous in R.*

Proof. Let $\Gamma^n = \{\Gamma_\alpha^n\}$ and let $\mathfrak{V}_1$ be a covering of $R\times I$ in which Γ_1^n is defined. We define a P-covering $\mathfrak{V}_2$ which is a refinement of $\mathfrak{V}_1$. Each point $(x,t)\in R\times I$ is contained in a set $U\times W$ such that U is an open set R, W is an interval (open in I) of I, and $U\times W$ is contained in some open set of $\mathfrak{V}_1$. The compact set $x\times I$ is covered by a finite number of these product sets $U_i(x)\times W_i(x)$. Let $\Pi_i U_i(x) = U(x)$. Then the sets $U(x)\times W_i(x)$ cover $x\times I$. It may be assumed that none of the intervals $W_i(x)$ is superfluous, i.e. that $\Sigma_{j\neq i}U_j(x) \neq I$. It may also be assumed that the sets $W_i(x)$ are divided if necessary so that no $W_i(x)\supset I$. The sets $U(x)$ form a covering of R. Since R is paracompact this covering has a locally finite refinement $\mathfrak{U} = \{U_\alpha\}$. Coresponding to each U_α choose x so

that $U_a \subset U(x)$ and let $W_{ai} = W_i(x)$. Then the sets $V_{ai} = U_a \times W_{ai}$ form a P-covering $\mathfrak{V}_2$ of $R \times I$ and $\mathfrak{V}_2$ is a refinement of $\mathfrak{V}_1$.

We may assume that $\Gamma_2{}^n \varepsilon \Gamma^n$ is defined in the nerve N_2 of $\mathfrak{V}_2$. Let $\mathfrak{U}_2 = \{U_a\}$ and let $\mathfrak{U}'_2$ and $\mathfrak{U}''_2$ be the corresponding coverings $\{U_a \times 0\}$ and $\{U_a \times 1\}$ of $R \times 0$ and $R \times 1$. Then $N'_2 = N(\mathfrak{U}')$ and $N''_2 = N(\mathfrak{U}''_2)$ are subcomplexes of N_2. We may assume that $D_{0,2}{}^n \varepsilon D_0{}^n$ and $D_{1,2}{}^n \varepsilon D_1{}^n$ are the part of $\Gamma_2{}^n$ in N'_2 and N''_2. Hence by Lemma 6.2 the corresponding cocycles of $N(\mathfrak{U}_2)$ are cohomologous. But the cocycles corresponding to $D_{0,2}{}^n$ and $D_{1,2}{}^n$ are $C_{0,2}{}^n \varepsilon C_0{}^n$ and $C_{1,2}{}^n \varepsilon C_1{}^n$. Hence $C_{0,2}{}^n \sim C_{1,2}{}^n$ in $N(\mathfrak{U}_2)$. Hence $C_0{}^n \sim C_1{}^n$ in R.

7. Hopf's classification theorem. Let $K \times I$ be the topological product of the simplicial complex K by the line interval I. Then $K \times I$ can be considered as a complex having as cells the simplexes $\sigma_a{}^n \times 0$ and $\sigma_a{}^n \times 1$ and the cells $\sigma_a{}^n \times J$ where $\sigma_a{}^n$ is any simplex of K and J is the interior of the interval I. The closure of each cell of $K \times I$ is a finite complex and the boundary of the boundary of each cell is zero. Hence one can define chains, coboundaries and cocycles in $K \times I$. The complexes $K \times 0$ and $K \times 1$ are isomorphic to K under the correspondences $\sigma_a{}^n \leftrightarrow \sigma_a{}^n \times 0$ and $\sigma_a{}^n \leftrightarrow \sigma_a{}^n \times 1$. Thus to each chain C^n of K there correspond chains C'^n of $K \times 0$ and C''^n of $K \times 1$, and conversely.

LEMMA 7.1. *Let $C_0{}^n$ and $C_1{}^n$ be chains of an n-dimensional simplicial complex K and let $D_0{}^n$ and $D_1{}^n$ be the corresponding chains of $K \times 0$ and $K \times 1$. Then $D_0{}^n$ and $D_1{}^n$ are the parts in $K \times 0$ and $K \times 1$ of a cocycle Γ^n of $K \times I$ if and only if $C_0{}^n \sim C_1{}^n$, in K.*

Proof. Whitney's proof ([28], p. 53) of this lemma for the case of a finite complex carries over without change.

We now assume that S^n, $(n \geq 1)$, is a simplicial complex, the boundary of an $n+1$-simplex. We assume that the fundamental cocycle of S^n is $+\sigma_0{}^n$ where $\sigma_0{}^n$ is a designated simplex of S^n.

We write ψ_0 and ψ_1 for the trivial mappings of K on $K \times 0$ and $K \times 1$: for $p \varepsilon K$, $\psi_0(p) = (p, 0)$, $\psi_1(p) = (p, 1)$.

LEMMA 7.2. *Let g_0 and g_1 be two barycentric mappings of an n-dimensional simplicial complex K into S^n. If $g^*{}_0\Delta^n \sim g^*{}_1\Delta^n$ in K then there is a mapping h of $K \times I$ in S^n continuous on each finite subcomplex of $K \times I$ and such that $h \mid K \times 0 = g_0\psi_0{}^{-1}$, $h \mid K \times 1 = g_1\psi_1{}^{-1}$.*

Proof. Projecting from a point inside σ_0^n one deforms g_0 and g_1 into normal mappings and then Whitney's proof ([28], p. 55) applies.

LEMMA 7.3. *Let R be a normal space with $\dim R \leq n$. Let Z be the family of locally finite coverings or the star-finite coverings or the finite coverings of R. If f_0 and f_1 are two mappings of R into the n-sphere S^n such that $f^*_0\Delta^n \sim f^*_1\Delta^n$ in Z then f_0 and f_1 are homotopic.*

Proof. Let $f^*_0\Delta^n = \Gamma'^n = \{\Gamma'_\alpha{}^n\}$ and $f^*_1\Delta^n = \Gamma''^n = \{\Gamma''_\alpha{}^n\}$. Let $\mathfrak{V}$ be the covering of S^n by the stars of its vertices and let $\mathfrak{U}_0 = f_0^{-1}\mathfrak{V}$ and $\mathfrak{U}_1 = f_1^{-1}\mathfrak{V}$. Let $\mathfrak{U}_2$ be a common refinement in Z of $\mathfrak{U}_0$ and $\mathfrak{U}_1$. Let $\mathfrak{U}_3$ be a refinement in Z of $\mathfrak{U}_2$ such that $\Gamma'_3{}^n \sim \Gamma''_3{}^n$ in N_3. Let $\mathfrak{U}_4$ be a refinement in Z of $\mathfrak{U}_3$ such that $\dim N_4 \leq n$. Let N_0 and N_1 be identified with the corresponding subcomplexes of S^n and let $\pi_0{}^4$ and $\pi_1{}^4$ be projections from N_4 to N_0 and N_1 respectively. Then $\pi_0{}^4$ and $\pi_1{}^4$ are simplicial mappings of N_4 into S^n and $\pi^*_0{}^4\Delta^n \sim \Gamma'_4{}^n \sim \Gamma''_4{}^n \sim \pi^*_1{}^4\Delta^n$ in N_4. Hence by Lemma 7.2 the corresponding mappings $\pi_0{}^4\psi_0^{-1}$ of $N_4 \times 0$ and $\pi_1{}^4\psi_1^{-1}$ of $N_4 \times 1$ can be extended to a mapping h of $N_4 \times I$ into S^n, with h continuous on each finite subcomplex of $N_4 \times I$. Let ϕ be a canonical mapping of R in N_4. Then ϕ maps a neighborhood of each point $p \in R$ into a finite subcomplex K of N_4 and h maps $K \times I$ continuously in S^n. Hence $h\phi$ is continuous and hence the mappings $\pi_0{}^4\phi$ and $\pi_1{}^4\phi$ of R in S^n are homotopic. Any point $p \in R$ is mapped by ϕ into the closure of the simplex $\sigma_4(p)$ determined by p in N_4 and this is mapped by $\pi_0{}^4$ into the closure of the simplex determined by p in N_0, a closed simplex which also contains $f_0(p)$. Hence $\pi_0{}^4\phi$ is uniformly homotopic to f_0. Similarly $\pi_1{}^4\phi$ is uniformly homotopic to f_1. Therefore f_0 is homotopic to f_1, which completes the proof.

LEMMA 7.4. *Let R be a normal space with $\dim R \leq n$. Let Z be the family of locally finite coverings or the star-finite coverings or the finite coverings of R. Then, from the correspondence between the cocycle $f^*\Delta^n$ and the mapping f, each element of the n-cohomology group of R in Z corresponds to one and only one class of homotopic mappings of R into the n-sphere S^n.*

Proof. By Lemma 7.3 if two mappings correspond to cocycles of the same cohomology element they are homotopic and hence each cohomology element corresponds to at most one homotopy class. Hence it is sufficient to show that to each cohomology element corresponds some mapping. Thus we must show that if Γ^n is an n-cocycle of Z there is a mapping f of R in S^n such that $f^*\Delta^n \sim \Gamma^n$ in Z.

Let $\Gamma^n = \{\Gamma_a{}^n\}$ be any n-cocycle of Z and let $\mathfrak{U}_1$ be a normal covering of R in Z such that $\Gamma_1{}^n$ is defined in the at most n-dimensional nerve N_1 of $\mathfrak{U}_1$. Let g_1 be a mapping of N_1 into S^n which maps the $n-1$ dimensional part $N_1{}^{n-1}$ of N_1 into a designated vertex p_0 of S^n and which maps each n-simplex $\sigma_a{}^n$ of N_1 so that, for some simplicial subdivision $\overline{\sigma}'_a{}^n$ of $\overline{\sigma}_a{}^n$ in which the boundary $\delta\sigma_a{}^n$ is not subdivided, $\overline{\sigma}'_a{}^n$ is mapped simplicially and exactly $|a^a|$ n-simplexes of $\sigma'_a{}^n$ are mapped on a designated n-simplex $\sigma_0{}^n$ of S^n, the mapping preserving orientation on these simplexes if a^a is positive, otherwise reversing orientation. The construction of such a subdivision and mapping is easy and omitted. Let N_2 be the resulting subdivision of N_1 and g_2 the resulting simplicial mapping of N_2 in S^n. Let ϕ_2 be the mapping of R on N_2 corresponding to a canonical mapping ϕ_1 of R on N_1. Let f be the mapping $g_1\phi_1 = g_2\phi_2$ of R in S^n. Let $f^*\Delta^n = C^n$ in Z.

If $\mathfrak{B}_1$ is the covering of S^n by the stars of its vertices, the complex S^n is the nerve of $\mathfrak{B}_1$. We assume that the cocycle of Δ^n on this nerve is $\Delta_1{}^n = +\sigma_0{}^n$. Let $\mathfrak{U}_3 = f^{-1}(V_1)$. Its nerve N_3 is S^n or a subcomplex of S^n. Let ψ be the identical mapping of N_3 in S^n. Let $\mathfrak{U}_2$ be the inverse image by ϕ_2 of the covering of N_2 by the stars of its vertices. Then $\mathfrak{U}_2$ is in Z, its nerve is N_2 and it is a common refinement of $\mathfrak{U}_1$ and $\mathfrak{U}_3$. The mapping $\pi_3{}^2 = \psi^{-1}g_2$ is a projection from N_2 to N_3. Let $\pi_1{}^2$ be a projection from N_2 to N_1 which maps exactly one simplex of $\sigma'_a{}^n$ into $\sigma_a{}^n$. Then $\pi^*{}_1{}^2\Gamma_1{}^2$ has coefficient a^a on one simplex of $\sigma'_a{}^n$ and $\pi^*{}_3{}^2C_3{}^n$ has coefficient $\operatorname{sgn} a^a$ on $|a^a|$ simplexes of $\sigma'_a{}^n$. Hence $\pi^*{}_3{}^2C_3{}^n \sim \pi^*{}_1{}^2\Gamma_1{}^n$ in N_2. Hence $C^n \sim \Gamma^n$ in Z. Therefore $f^*\Delta^n \sim \Gamma^n$ in Z, which completes the proof.

THEOREM 7.5. (*Hopf classification theorem*). *Let R be a paracompact normal space with* $\dim R \leq n$, $(n \geq 1)$. *Then the correspondence between the cocycle $f^*\Delta^n$ and the mapping f induces a $1-1$ correspondence between the elements of the n-cohomology group of R and the classes of homotopic mappings of R into the n-sphere S^n.*

Proof. The family Z of locally finite coverings of R forms a cofinal family and hence each element of $H^n(R)$ determines a unique element of $H_Z{}^n(R)$ which by Lemma 7.4 corresponds to a unique homotopy class of mappings of R in S^n.

Conversely, let f_0 and f_1 be elements of the same homotopy class of mappings of R in S^n. Let h be a mapping of $R \times I$ in S^n such that for $x \in R$, $h(x,0) = f_0(x)$, $h(x,1) = f_1(x)$. Then $f^*{}_0\Delta^n$ and $f^*{}_1\Delta^n$ are the cocycles

of R corresponding to the parts in $R \times 0$ and $R \times 1$ of the cocycle $h^*\Delta^n$ in $R \times I$. Hence, by Lemma 6.3, ${f^*}_0\Delta^n \sim {f^*}_1\Delta^n$ in R. Thus, to each homotopy class of mappings of R in S corresponds a unique element of $H^n(R)$. This completes the proof.

8. Bruschlinsky's theorem. In the special case of mappings into a circumference S^1, restrictions on the dimension of the space R can be removed. This is a consequence of the fact that every mapping of an n-sphere, $n > 1$, in a circumference S^1 is inessential. We briefly indicate the changes necessary to prove the more general mapping theorem when the image sphere is S^1.

a) Let g_0 and g_1 be two mappings of a simplicial complex K into the circumference S^1, each continuous on each finite subcomplex of K, such that all vertices of K are mapped by g_0 and g_1 into a designated point p_0 of S^1. If the degrees of the mappings g_0 and g_1 for the simplex ${\sigma_\alpha}^1$ of K are a^α and b^α respectively and if ${C_0}^1 = a^\alpha {\sigma_\alpha}^1$ and ${C_1}^1 = b^\alpha {\sigma_\alpha}^1$ are cohomologous cocycles of K, then there is a mapping h of $K \times I$ in S^1, continuous on each finite subcomplex of $K \times I$, and such that, if $x \in K$, $h(x, 0) = g_0(x)$ and $h(x, 1) = g_1(x)$.

By Lemma 7.1 there is a cocycle $\Gamma^1 = a^\alpha {\sigma'^1}_\alpha + b^\alpha {\sigma''^1}_\alpha + c^\beta(\sigma_\beta{}^0 \times I)$ of $K \times I$ whose parts on $K \times 0$ and $K \times 1$ correspond to ${C_0}^1$ and ${C_1}^1$. Let the partial mapping by h of $K \times 0$ and $K \times 1$ be defined by $h(x, 0) = g_0(x)$ and $h(x, 1) = g_1(x)$, and let the mapping be extended to the remaining 1-cells of $K \times I$ by mapping $\sigma_\beta{}^0 \times I$ with degree c^β. Since Γ^1 is a cocycle, the boundary of each 2-cell ${\sigma_\alpha}^1 \times I$ is mapped with degree zero, and the mapping can be extended over the 2-cells. The boundary of each 3-cell is now mapped in S^1, necessarily inessentially, and hence the mapping can be extended to all 3-cells. And so on. The mapping h is thus defined on all cells of $K \times I$.

b) Let g_0 and g_1 be two simplicial mappings of a simplicial complex K in the boundary S^1 of a 2-simplex. If ${g^*}_0\Delta^1 \sim {g^*}_1\Delta^1$ in K there is a mapping h of $K \times I$ in S^1, continuous on each finite subcomplex of $K \times I$, and such that, if $x \in K$, $h(x, 0) = g_0(x)$ and $h(x, 1) = g_1(x)$.

We assume that $\Delta^1 = +\sigma_0{}^1$ where $\sigma_0{}^1$ is a designated simplex of S^1. Let p_0 be a point of S^1 not in $\sigma_0{}^1$. Projecting from a point of $\sigma_0{}^1$ one deforms the mappings g_0 and g_1 so that all vertices are mapped into p_0. Thus, b) is seen to be a consequence of a).

c) Let R be a normal space and let Z be the family of locally finite coverings or the star-finite coverings or the finite coverings of R. If f_0 and f_1

are mappings of R in the circumference S^1 such that $f^*_0\Delta^1 \sim f^*_1\Delta^1$ in Z then f_0 and f_1 are homotopic.

This follows from b) using the method of proof of Lemma 7.3.

d) If Γ^1 is a 1-cocycle of R in Z there is a mapping f of R in S^1 such that $f^*\Delta^1 \sim \Gamma^1$ in Z.

Let $\mathfrak{U}_1$ be a covering of R in which $\Gamma_1{}^1 = a^\alpha \sigma_\alpha{}^1$ is defined. Let ϕ_1 be a canonical mapping of R in the nerve N_1 of $\mathfrak{U}_1$. Let g_1 be the following mapping of N_1 in S^1. Let each vertex of N_1 be mapped in a designated vertex p_0. Let each 1-simplex $\sigma_\alpha{}^1$ of N_1 be subdivided so it can be mapped barycentrically in S^1 with degree a^α. We can assume that exactly $|a^\alpha|$ simplexes of the subdivision $\sigma'_\alpha{}^1$ of $\sigma_\alpha{}^1$ are mapped in a designated 1-simplex $\sigma_0{}^1$ of S^1. Since $\Gamma_1{}^1$ is a cocycle the boundary of each 2-simplex of N_1 is mapped with degree zero. Hence the mapping can be extended over each 2-simplex. This extended mapping can be deformed into a simplicial mapping of a subdivision $\sigma'_\beta{}^2$ of $\sigma_\beta{}^2$ in such a way that the subdivisions $\sigma'_\alpha{}^1$ of simplexes $\sigma_\alpha{}^1$ of the boundary and the mappings of the boundary simplexes are unchanged. Since each boundary of a 3-simplex is mapped inessentially in S^1 the mapping can be extended to each 3-simplex and deformed into a suitable simplicial mapping, and so on. Thus we get a simplicial mapping g_2 of a subdivision N_2 of N_1 into S^1. If ϕ_2 is the mapping of R in N_2 corresponding to the mapping ϕ_1 in N_1 then $f = g_1\phi_1 = g_2\phi_2$ is a continuous mapping of R in S^1. As in the proof of Lemma 7.4 one proves that $f^*\Delta^1 \sim \Gamma^1$ in Z.

e) If R is a paracompact normal space, the correspondence between the cocycle $f^*\Delta^1$ and the mapping f induces a $1-1$ correspondence between the elements of the 1-cohomology group of R and the classes of homotopic mappings of R into the circumference S^1.

This follows from d) and Lemma 6.3.

The mappings into a circumference form a group. Thus, if we identify S^1 with the additive group of real numbers mod 1, we define the sum $f_0 + f_1$ of the mappings f_0 and f_1 of R in S^1 by $(f_0 + f_1)(p) = f_0(p) + f_1(p) \mod 1$ for each $p \in R$. The inessential mappings of R in S form a subgroup. The difference group of the mappings of R in S^1 mod the inessential mappings is a group whose elements are the classes of homotopic mappings of R in S^1.

Let Γ^1, C^1, D^1 be 1-cocycles of a paracompact normal space R such that $\Gamma^1 \sim C^1 + D^1$. Let U_1 be a locally finite covering of R such that $\Gamma_1{}^1 \sim C_1{}^1 + D_1{}^1$ in N_1. We may assume in fact that $\Gamma_1{}^1 = C_1{}^1 + D_1{}^1$ in N_1. Let $\Gamma_1{}^1 = a^\alpha \sigma_\alpha{}^1$, $C_1{}^1 = c^\alpha \sigma_\alpha{}^1$, $D_1{}^1 = d^\alpha \sigma_\alpha{}^1$. Then $a^\alpha = c^\alpha + d^\alpha$. Let g_0, g_1, g_2

be simplicial mappings of subdivisions of N_1 in S^1 mapping the vertices of N_1 into a vertex p_0 of S^1 and mapping subdivisions of $\sigma_a{}^1$ in S^1 with degrees a^a, c^a and d^a respectively. Then $g_1 + g_2$ maps $\sigma_a{}^1$ in S^1 with the same degree $a^a = c^a + d^a$ as g_0. Hence, by a), there is a mapping h of $N_1 \times I$ in S^1, continuous on each finite subcomplex, such that, if $x \in N_1$, $h(x, 0) = g_0(x)$ and $h(x, 1) = (g_1 + g_2)(x)$. Let ϕ be a canonical mapping of R in N_1. Let $f_0 = g_0\phi$, $f_1 = g_1\phi$, $f_2 = g_2\phi$. Then $f^*{}_0\Delta^1 \sim \Gamma^1$, $f^*{}_1\Delta^1 \sim C^1$, and $f^*{}_2\Delta^1 \sim D^1$. And $f_1 + f_2 = (g_1 + g_2)\phi$ is homotopic to $g_0\phi = f_0$. Hence, under the correspondence between $f^*\Delta^1$ and f, the sum of the cocycles corresponds to the sum of the corresponding mappings. Thus we have the following form of Bruschlinsky's Theorem:

THEOREM 8.1. (*Bruschlinsky's Theorem*). *If R is a paracompact normal space the correspondence between the cocycle $f^*\Delta^1$ and the mapping f induces an isomorphism between the* 1*-cohomology group of R and the group of homotopy classes of mappings of R into the circumference S^1.*

9. Finite coverings. A compact space is a paracompact space in which the family of finite coverings forms a cofinal family. Hence the Hopf theorem and Bruschlinsky's theorem hold for compact normal spaces even if cocycles and cohomologies are based on the family of finite coverings. Some similar theorems stated in terms of finite coverings hold, also, for non-compact spaces. We shall return to consideration of these theorems after a digression.

We first show that homotopic mappings of a countably compact space, or in particular a compact space, into a metric space are uniformly homotopic.

LEMMA[16] 9.1. *Let R be a countably compact space, T a compact space in which the first countability axiom is satisfied, and $\mathfrak{U}$ a countable or finite covering of $R \times T$ by open sets. Then there exist finite coverings $\mathfrak{V}$ and $\mathfrak{W}$ of R and T respectively such that $\mathfrak{V} \times \mathfrak{W}$ is a refinement of $\mathfrak{U}$.*

Proof. Since R is countably compact so is $R \times t$ and hence, because $R \times t$ is covered by the countable family $\mathfrak{U}$ of open sets, it is covered by a finite number of these, say $\{U_1(t), \cdots, U_r(t)\}$. We define in succession open sets $V_i{}^t$ of R and $W_i{}^t$ of T, such that $V_i{}^t \times W_i{}^t \subset U_i(t)$ and such that

$$\{V_1{}^t \times W_1{}^t, \cdots, V^t{}_{j-1} \times W^t{}_{j-1}, U_j(t), \cdots, U_r(t)\}$$

[16] A similar theorem was proved by Čech ([9], p. 3).

covers $R \times t$. Suppose that $V_i{}^t$, $W_i{}^t$ have been defined for $i < j$. We must define V_j and W_j so that $F_j{}^t \subset V_j{}^t \times W_j{}^t \subset U_j{}^t$ where $F_j{}^t$ is the closed set of $R \times t$: $F_j{}^t = R \times t - (R \times t) \cdot [V_1{}^t \times W_1{}^t + \cdots + V^t{}_{j-1} \times W^t{}_{j-1} + U_{j+1}(t) + \cdots + U_r(t)]$. For each point $(x, t) \in F_j{}^t$ there are open sets $V_j(x, t)$ of R and $W_j(x, t)$ of T such that $(x, t) \in V_j(x, t) \times W_j(x, t) \subset U_j(t)$. Let $\{N_t{}^n\}$ be a complete family of neighborhoods of t in T such that $N_t{}^n \supset N_t{}^{n+1}$. If $G_{jn}{}^t = \Sigma V_j(x, t)$, where the summation is over all x such that $W_j(x, t) \supset N_t{}^n$, then $G_{jn}{}^t$ is an open set of R, $G_{jn}{}^t \times N_n{}^t \subset U_j(t)$, and $G_{jn}{}^t \subset G^t{}_{j,n+1}$. Since every $W_j(x, t)$ contains some neighborhood $N_t{}^n$, therefore $\sum_{n=1}^{\infty} G_{jn}{}^t \times t \supset F_j{}^t$. Therefore, since $F_i{}^t$ is countably compact, there is an n such that $G_{jn}{}^t \times t \supset F_j{}^t$. Let this set $G_{jn}{}^t$ be called $V_j{}^t$ and let the corresponding neighborhood $N_t{}^n$ be called $W_j{}^t$. Then, as required, $F_j{}^t \subset V_j{}^t \times W_j{}^t \subset U_j(t)$.

Thus eventually $R \times t$ is covered by the sets $\{V_i{}^t \times W_i{}^t\}$, where $\mathfrak{V}(t) = \{V_i{}^t\}$ is a finite covering of R. Let $W(t) = \Pi_i W_i{}^t$. A finite number of the open sets $W(t)$ cover the compact space T. Let this finite covering of T be $\mathfrak{W} = \{W(t_k)\}$. Let $\mathfrak{V}$ be a common refinement of the finite coverings $\mathfrak{V}(t_k)$. Then $\mathfrak{V} \times \mathfrak{W}$ is a finite covering of $R \times T$. Each open set of the covering $\mathfrak{V} \times \mathfrak{W}$ is, for some t, of the form $V \times W(t)$ where V is contained in an open set $V_i{}^t$ of $\mathfrak{V}(t)$ and $V_i{}^t \times W(t) \subset V_i{}^t \times W_i{}^t \subset U_i(t)$ which is an open set of $\mathfrak{U}$. Thus $\mathfrak{V} \times \mathfrak{W}$ is a refinement of $\mathfrak{U}$. This completes the proof.

It is an immediate consequence that, if R and T satisfy the hypotheses of Lemma 9.1, $R \times T$ is countably compact.

LEMMA 9.2. *Two continuous mappings f_0 and f_1 of a countably compact space R in a metric space S are uniformly homotopic if and only if they are homotopic.*

Proof. It follows immediately from the definitions that if f_0 and f_1 are uniformly homotopic they are homotopic.

Let f_0 and f_1 be homotopic. Then there is a continuous mapping h of $R \times I$ in S such that $h(x, 0) = f_0(x)$, $h(x, 1) = f_1(x)$. Since R is countably compact, and I is compact and satisfies the first countability axiom, $R \times I$ is countably compact. Since $h(R \times I)$ is the continuous image of a countably compact space it is a countably compact subspace of S and hence a compact metric space. Hence, for any $\epsilon > 0$, $h(R \times I)$ has a finite covering by open sets, each of diameter $< \epsilon$. Their inverse images form a finite covering of

$R \times I$ by open sets. Let this covering be $\mathfrak{U}$. Then, by Lemma 9. 1, there are coverings $\mathfrak{V}$ and $\mathfrak{W}$ of R and I respectively such that $\mathfrak{V} \times \mathfrak{W}$ is a refinement of $\mathfrak{U}$, and hence such that $h(V \times W)$ has diameter $< \epsilon$ for any $V \in \mathfrak{V}$ and $W \in \mathfrak{W}$.

Since I is compact there is a number $\delta > 0$ such that for every pair of points t, t' of I such that $|t - t'| < \delta$, there is an open set $W \in \mathfrak{W}$ which contains both points. Hence, if $|t - t'| < \delta$, (x, t) and (x, t') are in the same set $V \times W$, and hence $\rho(h(x, t), h(x, t')) < \epsilon$. Therefore f_0 and f_1 are uniformly homotopic, which completes the proof.

We now recall some properties of the Tychonoff-Čech compaction of a normal space. Čech [10] has shown that if R is a normal space there is a compact Hausdorff space $\beta(R)$ and a mapping ψ of R into $\beta(R)$ such that:

1) The image set ψR is a normal Hausdorff space which is dense in $\beta(R)$.

2) Any bounded continuous real function f on R has the form $f(x) = F\psi(x)$ where F is a similar function on $\beta(R)$. By 1) the function F is unique.

3) It follows from 2) that any mapping f of R into S^n determines a mapping F of $\beta(R)$ into S^n, such that $f(x) = F\psi(x)$.

We need these facts in the proof of

THEOREM 9. 3. *Let R be a normal space with* $\dim R \leq n$, $(n \geq 1)$. *The correspondence between the cocycle $f^*\Delta^n$ and the mapping f induces a $1-1$ correspondence between the elements of the n-cohomology group $H_F^n(R)$ based on finite coverings and the classes of uniformly homotopic mappings of R into the n-sphere S^n.*

Proof. Corresponding to each mapping f of R in S^n is a mapping F of $\beta(R)$ in S^n such that $f = F\psi$. This mapping f determines a cocycle $f^*\Delta^n = \psi^* F^* \Delta^n$. Thus, corresponding to a mapping f is a mapping F and the cocycle $F^*\Delta^n$ as well as the cocycle $f^*\Delta^n$. Since $\beta(R)$ is compact, and hence paracompact, the correspondence between F and $F^*\Delta^n$ leads to a $1-1$ correspondence between the homotopy classes of mappings of $\beta(R)$ in S^n and the elements of the n-cohomology group of $\beta(R)$. It is sufficient to show that a) the correspondence between f and F leads to a $1-1$ correspondence between the uniform homotopy classes of mappings of R in S^n and the homotopy classes of mappings of $\beta(R)$ in S^n, and that b) the mapping ψ^* of the cocycles of $\beta(R)$ into the cocycles of R induces an isomorphism between the cohomology group $H^n(\beta(R))$ and the cohomology group $H_F^n(R)$.

a) First let F_0 and F_1 be homotopic mappings of $\beta(R)$ in S^n, and let f_i map R in S^n so that $f_i(x) = F_i\psi(x)$. Since $\beta(R)$ is compact, F_0 and F_1 are uniformly homotopic (by Lemma 9.2) and hence the same is true of f_0 and f_1.

On the other hand let f_0 and f_1 be uniformly homotopic mappings of R in S^n and let F_0, F_1 be the corresponding mappings of $\beta(R)$. Let $g(x,t)$ be the mapping of $R \times I$ in S^n which gives the uniform homotopy of f_0 and f_1. For fixed t, $g(x,t)$ is a continuous mapping of R in S^n which we call $g_t(x)$ and there is a corresponding continuous mapping G_t of $\beta(R)$ in S^n such that $G_t\psi(x) = g_t(x)$. Defining $G(y,t)$ to be $G_t(y)$ for $y \in \beta(R)$ we have a transformation of $\beta(R) \times I$ into S^n with $G(y,0) = F_0(y)$; $G(y,1) = F_1(y)$. Hence, if we can show that $G(y,t)$ is continuous, F_0 and F_1 are homotopic.

Let $\epsilon > 0$, and let $(y_0, t_0) \in \beta(R) \times I$. Since G_{t_0} is continuous there is a neighborhood U of y_0 in $\beta(R)$ such that, for $y' \in U$, the distance $\rho(G_{t_0}(y_0), G_{t_0}(y')) < \epsilon/3$. Since $g(x,t)$ is a uniform homotopy there is a δ-neighborhood W of t_0 such that, if $x \in R$ and $t \in W$, $\rho(g_{t_0}(x), g_t(x)) = \rho(G_{t_0}\psi(x), G_t\psi(x)) < \epsilon/3$. Now let (y,t) be any point of the neighborhood $U \times W$ of (y_0, t_0) in $\beta(R) \times I$. Since G_t is continuous there is a neighborhood $U_1 \subset U$ of y such that, if $y' \in U_1$, $\rho(G_t(y'), G_t(y)) < \epsilon/3$.

But since ψR is dense in $\beta(R)$ there is a point $x \in R$ such that $\psi(x) \in U_1$. Hence $\rho(G(y_0, t_0), G(y,t)) \leq \rho(G_{t_0}(y_0), G_{t_0}\psi(x)) + \rho(G_{t_0}\psi(x), G_t\psi(x)) + \rho(G_t\psi(x), G_t(y)) < \epsilon$. Hence F_0 and F_1 are homotopic. Hence we have a $1-1$ correspondence between the uniform homotopy classes of mappings of R in S^n and the homotopy classes of mappings of $\beta(R)$ in S^n.

b) We have still to show that the homomorphism of the cohomology groups induced by ψ^* is an isomorphism. Let $\mathfrak{U}$ be any finite covering of R on whose nerve a cocycle Γ^n of R is defined, or on whose nerve a cohomology $\Gamma^n \sim \Gamma'^n$ holds. Let $U_1, \cdots, U_k$ be the sets of $\mathfrak{U}$. As in Theorem 1.1 we can define continuous functions (barycentric coordinates) f_i on R such that, if $x \in R - U_i$, $f_i(x) = 0$ and, if $x \in R$, $\Sigma_i f_i(x) = 1$. Let F_i be the continuous functions on $\beta(R)$ such that $f_i(x) = F_i\psi(x)$ for $x \in R$. Then $F_i(y) = 0$ if $y \in \psi(R - U_i)$ and hence by continuity also if $y \in \overline{\psi(R - U_i)}$. Also $\Sigma F_i(y) = 1$ if $y \in \psi R$, and hence by continuity also if $y \in \overline{\psi R} = \beta(R)$. Hence, if $y \in \beta(R)$, for some i, $F_i(y) \neq 0$ and $y \in \beta(R) - \overline{\psi(R - U_i)}$. Hence the open sets $V_i = \beta(R) - \overline{\psi(R - U_i)}$ form a covering $\mathfrak{V}$ of $\beta(R)$. Hence the open sets $W_i = \psi^{-1}V_i$ form a covering $\mathfrak{W}$ of R. Now $\psi(W_i) \subset V_i = \beta(R) - \overline{\psi(R - U_i)} \subset \beta(R) - \psi(R - U_i)$. Hence W_i and $R - U_i$ do not intersect. Hence $W_i \subset U_i$ and $\mathfrak{W}$ is a refinement of $\mathfrak{U}$.

If any collection of open sets of $\mathfrak{W}$ intersect in a point $p \in R$ then $\psi(p)$ is in each of the corresponding sets of $\mathfrak{V}$ which therefore also intersect.

Conversely, if any collection of open sets of $\mathfrak{V}$ intersect, their intersection contains points of ψR and hence the corresponding sets of $\mathfrak{W}$ intersect. Hence the mapping of $N(\mathfrak{W})$ in $N(\mathfrak{V})$ induced by ψ is an isomorphism. Hence any cocycle or cohomology in the nerve of the refinement $\mathfrak{W}$ of $\mathfrak{U}$ is the image of a cocycle or cohomology in the nerve of a covering $\mathfrak{V}$ of $\beta(R)$. Hence the mapping ψ^* induces an isomorphism of $H^n(\beta(R))$ on $H_F^n(R)$. This completes the proof.

COROLLARY 9.4. *Let R be a countably compact normal space with* $\dim R \leq n$, $(n \geq 1)$. *The correspondence between the cocycle $f^*\Delta^n$ and the mapping f induces a $1-1$ correspondence between the elements of the n-cohomology group $H_F^n(R)$ based on finite coverings and the classes of homotopic mappings of R into the n-sphere S^n.*

Proof. Since by Lemma 9.2 the classes of homotopic mappings of the countably compact space R are identical with the classes of uniformly homotopic mappings, the corollary follows immediately from Theorem 9.3.

THEOREM 9.5. *If R is a normal* [*countably compact normal*] *space, the correspondence between the cocycle $f^*\Delta^1$ and the mapping f induces an isomorphism between the* 1*-cohomology group $H_F^1(R)$ based on finite coverings and the group of uniform homotopy* [*homotopy*] *classes of mappings of R into the circumference S^1.*

Proof. The proof of these forms of Bruschlinsky's Theorem is adequately indicated by the proofs of Theorem 9.3 and Corollary 9.4.

Theorem 9.5 may be used to find the 1-dimensional cohomology groups, based on finite coverings, of spaces for which the group of uniform homotopy classes of mappings in S^1 can be calculated. It is difficult to find these cohomology groups directly, even in the simplest non-compact spaces. In Theorem 9.6 the 1-cohomology group of the straight line is described. The remarkable results indicate that the cohomology groups based on finite coverings are not suitable for describing the connectivity properties of non-compact spaces.

THEOREM 9.6. *The* 1*-cohomology group of the real line based on finite coverings is isomorphic with the difference group of the additive group of all real continuous functions on the real line modulo the group of all bounded continuous functions. The number of its elements is $\aleph = 2^{\aleph_0}$.*

Proof. Let R^1 be the real line and let f be any mapping of R^1 in the unit circle S^1 in the complex plane. Since every such mapping is homotopic to a constant mapping there is a continuous real function ϕ on R^1 such that,[17] for $x \in R^1$, $f(x) = \exp(i\phi(x))$. To every such function ϕ corresponds a unique mapping f and to the sum or difference of such real functions ϕ corresponds the sum or difference of the mappings. Thus there is a homomorphism of the additive group of real continuous functions onto the group of mappings of R^1 in S^1. The bounded real continuous functions, and only these,[18] correspond to mappings of R^1 in S^1, which are uniformly homotopic to constant mappings. Hence the difference group of the real continuous functions modulo the bounded functions is isomorphic with the group of uniform homotopy classes of mappings of R^1 in S^1, which by Theorem 9.5 is isomorphic with the 1-cohomology group of R^1 based on finite coverings.

The power of the set of continuous real functions is $\aleph$. Also the set of functions $\phi(x) = kx$ with k real, which is in $1-1$ correspondence with a subset of the set of elements of the above difference group, has power $\aleph$. Hence the number of elements in the 1-cohomology group of R^1 based on finite coverings is $\aleph$.

10. Applications to non-algebraic topology. We now apply some of the preceding methods and results to the proof of two theorems of non-algebraic topology. The first is a variation of Borsuk's Theorem in which uniform homotopy is replaced by ordinary homotopy. The second is a generalization of a theorem of Alexandroff ([2], p. 220) on transforming a given space to a cyclic space of the same dimension by contracting a closed set to a point.

The following lemma is needed in the proof of Borsuk's Theorem with ordinary homotopy.

LEMMA 10.1. *Let R' be a closed subset of a paracompact normal space R and let g be any continuous real function on $R' \times I$. Then for any $\epsilon > 0$ there is a continuous real function h on $R \times I$ such that, for all $(x, t) \in R' \times I$, $|h(x, t) - g(x, t)| < \epsilon$.*

Proof. If R satisfies the Hausdorff separation axiom then, according to a theorem of Dieudonné ([11], Theorem 5), $R \times I$ is a paracompact

[17] See [13], p. 68.
[18] See [13], p. 68.

Hausdorff space and hence normal. A continuous function g on a closed subset $R' \times I$ of the normal space $R \times I$ can then be extended over the normal space and the lemma follows.

Otherwise, let $\mathfrak{V}$ be the covering of $R' \times I$ by the inverse images by g of a covering of the real line by intervals of length $< \epsilon$. Let $\mathfrak{V}_1 = \{V_{ai}\}$ be a P-covering of $R' \times I$ which is a refinement of $\mathfrak{V}$ (See proof of Lemma 6.3), and let $V_{ai} = U'_a \times W_{ai}$ where $\mathfrak{U}' = \{U'_a\}$ is a locally finite covering of R'. Let U'_a be extended to an open set U_a of R, such that $U'_a = U_a \cdot R'$, and let $U_0 = R - R'$. Then $\mathfrak{U} = \{U_a\}$, $\alpha = 0, 1, \cdots$, is a covering of R. Since R is paracompact and normal, there is a canonical mapping ϕ of R in the nerve N of $\mathfrak{U}$.

Let a point p_a be chosen in each set $U'_a \in \mathfrak{U}'$. Let G_t be the mapping of N in the real line which maps each vertex u_a of N, $\alpha \neq 0$, in $g(p_a, t)$ and maps u_0 in a fixed point, say the zero point, of the line, and maps N linearly. If $y \in N$, let $G(y, t) = G_t(y)$. Then G is continuous on each finite subcomplex of $N \times I$. If $x \in R$, let $h(x, t) = G(\phi(x), t)$. Then h is a continuous mapping of $R \times I$ in the real line.

Let $x \in R'$. Since $\mathfrak{U}'$ is locally finite, x is in a finite number of open sets of $\mathfrak{U}'$ and hence also is in a finite number of sets $U_a, \cdots, U_\gamma$ of $\mathfrak{U}$. Then ϕ maps x in the closure of the simplex $u_a \cdots u_\gamma$ determined by x in N, a simplex which does not have u_0 as a vertex. If $x \in U_a$, then $x \in U'_a$ and hence also, for each t, $(x, t) \in U'_a \times W_{ai}$ for some i. Also $p_a \in U'_a$ and $(p_a, t) \in U'_a \times W_{ai}$. Hence (x, t) and (p_a, t) are in the same set V_{ai} of the covering $\mathfrak{V}_1$ and hence also in the same set of the covering $\mathfrak{V}$. Hence $g(x, t)$ and $g(p_a, t)$ are in the same interval of the ϵ-covering of the real line. Hence, since $G_t(u_a) = g(p_a, t)$, G_t maps each vertex, and hence each point, of the closed simplex determined by x in a point of the real line within ϵ of $g(x, t)$. But one of the points of this simplex is $\phi(x)$. Hence, since $G_t\phi(x) = h(x, t)$, $|h(x, t) - g(x, t)| < \epsilon$. This completes the proof of the lemma.

Now, for normal spaces which are paracompact, we are able to prove Borsuk's Theorem with ordinary homotopy replacing uniform homotopy.

THEOREM [19] 10.2. *Let R' be a closed subset of a paracompact normal space R and let f_0 and f_1 be two homotopic mappings of R' into an absolute neighborhood retract S. If f_0 can be extended to a mapping F_0 of R into S, then f_1 can be similarly extended to F_1 with F_0 and F_1 homotopic.*

---

[19] This theorem, for a more restricted class of spaces, appears in my dissertation, Princeton University (1938). The proof was first published in [20], p. 86. For compact spaces the theorem is equivalent to Borsuk's theorem.

Proof. We assume that S is a subset of the Hilbert cube I_ω. Let $\eta > 0$ be such that the closed η-neighborhood $\bar{U}$ of S can be retracted onto S.

Let g be a mapping of $R' \times I$ in S such that, for $x \in R'$, $g(x,0) = f_0(x)$, $g(x,1) = f_1(x)$. Then it easily follows from Lemma 10.1 that there is a mapping h of $R \times I$ in I_ω such that, for $(x,t) \in R' \times I$, $\rho(g(x,t), h(x,t)) < \eta$. Then $h(R' \times I) \subset U$. Let $h_0(x,t)$ be the point dividing the segment $[F_0(x), h(x,0)]$ in the ratio $t : 1-t$. Then h_0 is a continuous mapping of $R \times I$ in I_ω such that $h_0(x,0) = F_0(x)$, $h_0(x,1) = h(x,0)$, and such that $h_0(R' \times I) \subset U$. Since R is normal, the mapping f_1 of R' in I_ω can be extended to a mapping F'_1 of R in I_ω. Then there is a continuous mapping h_1 of $R \times I$ in I_ω such that $h_1(x,0) = h(x,1)$, $h_1(x,1) = F'_1(x)$ and such that $h_1(R' \times I) \subset U$. Let h_2 be the continuous mapping of $R \times I$ in I_ω defined by $h_2(x,t) = h_0(x,3t)$ if $0 \leq t \leq 1/3$, $h_2(x,t) = h(x,3t-1)$ if $1/3 \leq t \leq 2/3$, and $h_2(x,t) = h_1(x,3t-2)$ if $2/3 \leq t \leq 1$. Then $h_2(x,0) = F_0(x)$, $h_2(x,1) = F'_1(x)$, and $h_2(R' \times I) \subset U$.

Let $h_2^{-1}U = V$. Then V is a neighborhood of $R' \times I$ in $R \times I$. Since I is compact, every point $x \in R'$ has a neighborhood $W(x)$ in R such that $W(x) \times I \subset V$. Let W be the sum of all these neighborhoods. When W is an open set of R such that $R' \subset W$, $W \times I \subset V$ and hence $h_2(W \times I) \subset U$. Since R is normal there is a continuous real function $p(x)$ defined on R such that $0 \leq p(x) \leq 1$ and such that $p(x) = 1$ for $x \in R'$ and $p(x) = 0$ for $x \in R - W$. Let h_3 be the continuous mapping of $R \times I$ in I_ω defined by $h_3(x,t) = h_2(x, t \cdot p(x))$. Then $h_3(x,0) = F_0(x)$, and, if $x \in R'$, $h_3(x,1) = f_1(x)$. Also $h_3(R \times I) \subset U$. Let ϕ be a retraction of U on S and let $H = \phi h_3$. Then H is a mapping of $R \times I$ in S such that $H(x,0) = F_0(x)$, and, if $x \in R'$, $H(x,1) = f_1(x)$. Let $F_1(x) = H(x,1)$. Then F_1 is an extension of f_1 homotopic to F_0 on S. This completes the proof of the theorem.

Let R' be a non-vacuous closed set of a space R. We say that a space R_1 is obtained from R by *contracting* R' to a point[20] if there is a continuous mapping ψ of R onto R_1 which maps R' into a closed set (p) consisting of a single point p of R_1 and which maps $R - R'$ homeomorphically onto $R_1 - (p)$.

If R is paracompact, R_1 is also paracompact. For let $\mathfrak{U} = \{U_\alpha\}$ be any covering of R_1 by open sets. Then $\{\psi^{-1}U_\alpha\}$ is a covering of R which, since R is paracompact, has a locally finite refinement $\mathfrak{V} = \{V_\beta\}$. Then $\{(R - R') \cdot V_\beta\}$ is a locally finite covering of $R - R'$ which, by the homeomorphism $\psi \mid R - R'$, is transformed into a locally finite covering $\{\psi((R - R') \cdot V_\beta)\}$ of $R_1 - (p)$ by open sets. By including one open set

[20] See [2], p. 220.

of $\mathfrak{U}$ containing p, we get a locally finite covering of R_1 which is a refinement of $\mathfrak{U}$. Hence R_1 is paracompact.

Similarly if R is compact R_1 is also compact. If R has a countable base for its open sets and if p has a countable base for its neighborhoods in R_1 then R_1 has a countable base for its open sets. For, since R has a countable base, $R - R'$ also has a countable base, and hence also $R_1 - (p)$ has a countable base. This together with the countable base for the neighborhoods of p forms a countable base for R_1.

If R has dimension $\leq n$, and if every open set of R_1 containing p contains the closure of an open set containing p then R_1 has dimension $\leq n$. For let $\mathfrak{U} = \{U_\alpha\}$ be any covering of R_1. If $p \in U_\gamma \in \mathfrak{U}$, let G_1, G_2 be neighborhoods of p such that $\bar{G}_1 \subset G_2$, $\bar{G}_2 \subset U_\gamma$. Then U_γ together with all the open sets $U_\alpha - \bar{G}_2$ form a covering $\mathfrak{U}_1$ of R_1. Then $\psi^{-1}\mathfrak{U}_1$ is a covering of R which has a refinement $\mathfrak{V}_1$ of order $\leq n+1$. All sets of $\mathfrak{V}_1$ which intersect $\psi^{-1}\bar{G}_2$ are contained in $\psi^{-1}U_\gamma \in \mathfrak{V}_1$. These sets may be replaced by a single set, their sum, which we call V_0, without increasing the order of the covering, and the resultant covering $\mathfrak{V}_2$ is still a refinement of $\psi^{-1}\mathfrak{U}_1$. Then all sets of $\mathfrak{V}_2$, except V_0, are open sets of $R - R'$ and their images by ψ are open sets of R_1. The set V_0 is the sum of the open sets $\psi^{-1}G_2$ and $V_0 - \psi^{-1}\bar{G}_1$. The image by ψ of $\psi^{-1}G_2$ is the open set G_2 of R_1 and the image of the open set $V_0 - \psi^{-1}\bar{G}_1$ of $R - R'$ is an open set of $R_1 - (p)$. Hence ψV_0 is an open set of R_1. Hence ψ maps the covering $\mathfrak{V}_2$ into an open covering $\mathfrak{U}_2$ of R_1. The order of $\mathfrak{U}_2$ is the same as that of $\mathfrak{V}_2$ and since $\mathfrak{V}_2$ is a refinement of $\mathfrak{V}_1$, $\mathfrak{U}_2$ is a refinement of $\mathfrak{U}_1$ and hence of $\mathfrak{U}$. Hence $\mathfrak{U}$ has a refinement of order $\leq n+1$. Hence R_1 has dimension $\leq n$.

If R is normal and if every open set of R_1 containing p contains the closure of an open set containing p, then R_1 is also normal. For let F_1 and F_2 be any two non-intersecting closed sets of R_1. The point p cannot be in both closed sets. Suppose that p is not in F_2. Let G_1 and G_2 be neighborhoods of p such that $\bar{G}_1 \subset G_2$ and $\bar{G}_2 \subset R_1 - F_2$. Then $R_1 - G_1$ is homeomorphic to a closed set of R and hence is normal. Hence there are non-intersecting open sets V_1 and V_2 of $R_1 - G_1$ containing the closed sets $(R_1 - G_1) \cdot F_1$ and F_2 respectively. Then $W_1 = G_2 + (R_1 - \bar{G}_1) \cdot V_1$ and $W_2 = (R - \bar{G}_2) \cdot V_2$ are non-intersecting open sets of R_1 which contain F_1 and F_2 respectively. Hence R_1 is normal.

An n-dimensional space is called cyclic if there is an essential mapping of R on the n-sphere S^n. The theorem of Alexandroff ([2] p. 220) [21] on

[21] See also [19], p. 195

obtaining a cyclic space by contracting a closed set to a point can be stated as follows for paracompact normal spaces.

THEOREM 10.3. *If R is a paracompact normal space of dimension n, $(n \leq 1)$, then, by contracting a suitable closed set R' of R to a point, R is transformed into a paracompact normal cyclic space R_1 of dimension n.*

Proof. By Corollary 3.5 there is a closed set R' of R and a mapping f of R' in the $(n-1)$-sphere S^{n-1} which cannot be extended to a mapping of R in S^{n-1}. Let S^{n-1} be the boundary of a simplex σ^n and let F be an extension of f to a mapping of R in $\overline{\sigma}^n$. Then $R' \subset F^{-1}S^{n-1}$. Let R' be replaced by a larger set, if necessary, so that $R' = F^{-1}S^{n-1}$. Let χ be a simplicial mapping of a subdivision of $\overline{\sigma}^n$ onto a sphere S^n such that S^{n-1} is mapped into a point q of S^n, with $\chi^{-1}(q) = S^{n-1}$, and such that the degree of the mapping is one.

We define a space R_1, a mapping ψ of R on R_1 and a mapping g of R_1 on S^n as follows: The space R_1 consists of a set of points in $1-1$ correspondence with the points of $R - R'$ and one additional point p. The mapping ψ maps each point of $R - R'$ on its corresponding point of R and maps R' into p. The mapping g maps each point $x \in R_1 - (p)$ into $\chi F \psi^{-1}(x)$ and maps p into q. Thus, for each point $y \in R$, $g\psi(y) = \chi F(y)$. Let the open sets of R_1 be the images by ψ of the open sets of $R - R'$ and the counter images by g of the open sets of S. Thus the mappings ψ and g are continuous, the set (p) is closed, and the partial mapping $\psi \mid R - R'$ is a homeomorphism. Hence R_1 is a space obtained from R by contracting R' to a point. The neighborhoods of q with radius $1/r$, where r is an integer, form a basis for the neighborhoods of q such that each contains the closure of another. The inverse images by g of these neighborhoods form a similar basis for the neighborhood of p in R. It follows that R_1 is a paracompact normal space of dimension $\leq n$ which is compact if R is compact and has a countable base if R has a countable base. Since also the intersection of the neighborhoods of p consists of p alone, R_1 can be shown to satisfy the T_0, T_1, T_2, or T_3 separation axioms if R does.

In proving the theorem we first take the case $n \geq 2$. The fundamental cocycle Δ^n of S^n is cohomologous in S^n to a cocycle of $S^n - (q)$. Replacing Δ^n, if necessary, by this cocycle we assume that Δ^n is a cocycle of $S^n - (q)$. Then $\chi^*\Delta^n$ is a cocycle of $\overline{\sigma}^n - S^{n-1}$. Since the mapping χ is of degree one, $\chi^*\Delta^n \sim \pm \Psi\Delta^{n-1}$ where $\Psi\Delta^{n-1}$ is the coboundary in $\overline{\sigma}^n - S^{n-1}$ of the fundamental cocycle Δ^{n-1} of S^{n-1}. We assume that $\chi^*\Delta^n \sim \Psi\Delta^{n-1}$. Then $F^*\chi^*\Delta^n \sim F^*\Psi\Delta^{n-1} \sim \Psi f^*\Delta^{n-1}$ in $R - R'$. But (See Theorem 5.2) $\Psi f^*\Delta^{n-1}$ and

hence also $F^*\chi^*\Delta^n$ is not ~ 0 in $R - R'$. Since $g\psi = \chi F$, $\psi^* g^*\Delta^n \sim F^*\chi^*\Delta^n$ in $R - R'$. Hence $\psi^* g^*\Delta^n$ is not ~ 0 in $R - R'$. Hence $g^*\Delta^n$ is not ~ 0 in $R_1 - (p)$. But every covering of R_1 has a refinement such that p is in only one open set of the refinement. Hence, since $n \geq 2$, $g^*\Delta^n$ is not ~ 0 in R_1. It follows that the dimension of R_1 is at least n, and hence is exactly n. It also follows, by Theorem 7.5, that g is an essential mapping of R_1 on S^n. Hence R_1 is a cyclic space of dimension n.

Now let $n = 1$. We assume that σ^1 is the line interval $(0, 1)$ and hence S^0 consist of the points 0 and 1. We also suppose that S^1 is the unit circle in the complex plane, that q is the unit point, and that $\chi(t) = \exp(2\pi i t)$ for $t \in \bar{\sigma}^1$. Suppose, if possible, that g is an inessential mapping of R_1 in S^1. Then, according to a theorem of Eilenberg ([13], p. 68) there is a real function ϕ of the points of R_1 such that $g(y) = \exp(2\pi i \phi(y))$ for $y \in R_1$. Since $g(y) = q$ only if $y = p$, $\phi(y)$ is an integer only if $y = p$. We may assume without loss of generality that $\phi(p) = 0$. Since $g\psi = \chi F$, if $x \in R$, $\exp(2\pi i F(x)) = \exp(2\pi i \phi\psi(x))$ and hence $\phi\psi(x) = F(x) \bmod 1$. Let A be the subset of R consisting of $F^{-1}(0) + \psi^{-1}\phi^{-1}(0, 1)$ where $(0, 1)$ is the open interval from 0 to 1. Using square brackets for closed intervals we can write, since $\phi\psi(x) = F(x) \bmod 1$, $A = F^{-1}[0, \frac{1}{2}] \cdot \psi^{-1}\phi^{-1}[0, \frac{1}{2}] + \psi^{-1}\phi^{-1}[\frac{1}{2}, 1]$. Thus A is a closed set. Also $A = F^{-1}[0, \frac{1}{2}) \cdot \psi^{-1}\phi^{-1}(-\frac{1}{4}, \frac{1}{2}) + \psi^{-1}\phi^{-1}(\frac{1}{4}, 1)$. Thus A is also an open set. Hence $B = R - A$ is a closed set. Now $f^{-1}(0) = F^{-1}(0) \subset A$ and $f^{-1}(1) = F^{-1}(1) \subset B$. Let F_1 be the mapping of R in S^0 which maps A in 0 and B in 1. Then F_1 is a continuous extension of f contrary to the assumption that f has no extension. Hence g is an essential mapping of R_1 in S^1. It follows that R_1 has dimension at least 1, and hence exactly 1. Hence R_1 is a cyclic space of dimension 1. This completes the proof of the theorem.

It may be noted that in the 1-dimensional case we have not used the assumption of paracompactness. It follows that any 1-dimensional normal space can be transformed, by contracting a closed set to a point, into a 1-dimensional normal cyclic space.

If in the proof of this theorem we used finite coverings and cocycles based on finite coverings and if we used Theorem 9.3 instead of Theorem 7.5 we could have shown that any n-dimensional normal space can be transformed, by contracting a closed set to a point, into an n-dimensional normal space which can be mapped essentially, in the sense of uniform homotopy, on the n-sphere.

11. Compact cocycles. It was shown by Borsuk and Eilenberg [6, 14] that the group of homotopy classes of mappings of the complement R of a

compact set of Euclidean space into the circumference S^1 is isomorphic with the cohomology group of R based on compact cocycles. This cohomology group is defined as follows:

Let $C_1, C_2, \cdots, C_a, \cdots$, be the compact subsets of a space R. A compact n-cocycle of R is a set $\{\Gamma_a{}^n\}$ of cocycles, one in each of the compact subsets $\{C_a\}$ such that, if $C_\beta \subset C_a$ and if $\psi_a{}^\beta$ is the identical mapping of C_β in C_a, then $\psi^*{}_a{}^\beta \Gamma_a{}^n \sim \Gamma_\beta{}^n$ in C_β. Two compact n-cocycles $\{\Gamma_a{}^n\}$ and $\{\Gamma'_a{}^n\}$ are cohomologous if $\Gamma_a{}^n \sim \Gamma'_a{}^n$ in C_a for every C_a. Also $\{\Gamma_a{}^n\} \sim \{\Gamma'_a{}^n\} - \{\Gamma''_a{}^n\}$ means $\Gamma_a{}^n \sim \Gamma'_a{}^n - \Gamma''_a{}^n$ in C_a for every C_a. The cohomology group of R based on compact cocycles is the group of classes of cohomologous compact cocycles of R.

If $\{D_\gamma\}$ is any subfamily of the compact subsets of R, a set $\{\Gamma_\gamma{}^n\}$ of cocycles, one on each of the sets $\{D_\gamma\}$, is called a cocycle of the subfamily if, whenever $D_\delta \subset D_\gamma$, if $\psi_\gamma{}^\delta$ is the identical mapping of D_δ in D_γ, $\psi^*{}_\gamma{}^\delta \Gamma_\gamma{}^n \sim \Gamma_\delta{}^n$ in D_δ. Similarly one defines the sum of two cocycles of the family or cohomology of two cocycles of the family. Then the classes of cohomologous cocycles of the family form the cohomology group of the family. A family $\{D_\gamma\}$ of compact subsets of R is called a cofinal family if every compact set of R is a subset of one of the sets $\{D_\gamma\}$. A cocycle of a cofinal family of compact sets determines a unique compact cocycle of R and the cohomology group of a complete family of compact sets is isomorphic with the cohomology group of R based on compact cocycles.

If Γ^n is an n-cocycle of R and if, for each compact set C_a of R, χ_a is the identical mapping of C_a in R, then $\chi^*{}_a \Gamma^n$ is an n-cocycle of C_a and $\{\chi^*{}_a \Gamma^n\}$ is a compact n-cocycle of R. If f is a mapping of R into S and if Γ^n is an n-cocycle of S, then $f^*\Gamma^n$ is a cocycle of R and $\{\chi^*{}_a f^* \Gamma^n\}$ is a compact cocycle of R. If $\Gamma^n \sim \Gamma'^n$ in S, then $\chi^*{}_a f^* \Gamma^n \sim \chi^*{}_a f^* \Gamma'^n$ in C_a, and $\{\chi^*{}_a f^* \Gamma^n\} \sim \{\chi^*{}_a f^* \Gamma'^n\}$ in R. If $\Gamma^n \sim \Gamma'^n - \Gamma''^n$ in S, then $\chi^*{}_a f^* \Gamma^n \sim \chi^*{}_a f^* \Gamma'^n - \chi^*{}_a f^* \Gamma''^n$ in C_a, and $\{\chi^*{}_a f^* \Gamma^n\} \sim \{\chi^*{}_a f^* \Gamma'^n\} - \{\chi^*{}_a f^* \Gamma''^n\}$ in R. Thus f induces a homomorphism of the n-cohomology group of S into the n-cohomology group of R based on compact cocycles.

Let R be a paracompact normal space which is locally compact and locally connected. The theorem of Borsuk and Eilenberg will be shown to hold for such a space.

Since R is locally connected, each component of R is an open set. For, if x is any point of R, any neighborhood of x contains a neighborhood which is contained in the component of R containing x. Hence the components V_a of R form an open covering $\mathfrak{V} = \{V_a\}$. Since R is locally compact, each point of R has a neighborhood whose closure is compact. Hence R has a covering

$\mathfrak{W} = \{W_a\}$ consisting of all open sets W_a with compact closures. Since R is paracompact, for any covering $\mathfrak{U}$ of R there is a common refinement $\mathfrak{U}_1$ of $\mathfrak{U}$, $\mathfrak{V}$, and $\mathfrak{W}$, which is locally finite. Since R is normal, we may further assume that $\mathfrak{U}_1$ is a normal covering of R, i. e., that there is a canonical mapping ϕ of R on the nerve N_1 of $\mathfrak{U}_1$ which is essential on every simplex of N_1. The star of each vertex u of N_1 is the image of an open set U of $\mathfrak{U}_1$, which is contained in an open set of $\mathfrak{W}$ whose closure is compact. Hence $\bar{U}$ is compact and its image, which is the closure of the star of u, is compact and hence is a finite complex. Thus $\mathfrak{U}_1$, is a star-finite covering and R is an s-space.

Since $\mathfrak{U}_1$ is a refinement of the covering $\mathfrak{V}$ by components, it follows that if U_α and U_β are sets of $\mathfrak{U}_1$ contained in distinct sets of $\mathfrak{V}$, i. e., in distinct components, the vertices u_α and u_β of N_1 are not connected. Thus ϕ does not map more than one component of R in any component of N_1. But the components of the complex N_1 are open and closed. Hence, since ϕ is continuous, the inverse image of a component of N_1 is an open and closed set of a component of R, and hence is a component of R.

Now, if L_1 is the star of a vertex of N_1, $A = \overline{\phi^{-1}(L_1)}$ is compact. Let L_2 be the star of the closed complex $\bar{L}_1$. Then $A_2 = \overline{\phi^{-1}(L_2)}$ is the sum of a finite number of compact sets and is therefore compact. Also A_1 is contained in the interior of A_2; $A_1 \subset\subset A_2$. Proceeding thus we get a sequence $\{L_i\}$ where L_{i+1} is the star of the finite complex $\bar{L}_i$ and a sequence of compact sets $\{A_i\}$ where $A_i = \overline{\phi^{-1}(L_i)}$ and $A_i \subset\subset A_{i+1}$. The sum ΣL_i of the sequence is seen to be a component of N_1. Hence ΣA_i is a component of R. Let p be a point of A_1. Let B_i be the component of A_i containing p. Then B_i is closed and therefore compact. Since R is locally connected, B_i is contained in an open set contained in a connected set of the interior of A_{i+1}. Hence $B_i \subset\subset B_{i+1}$. Also $\Sigma B_i = \Sigma A_i$. For otherwise there would be a point x of the connected set ΣA_i such that every neighborhood of x contains a point of ΣB_i and a point of $\Sigma A_i - \Sigma B_i$. But x is in the interior of some set A_i and has a neighborhood contained in a connected set of the interior of A_i. This neighborhood contains a point of some set B_i and a point of $A_i - \Sigma B_k$. Since the sets A_i are increasing it may be assumed that $i > j$. But then B_i would contain the whole neighborhood of x. Hence $\Sigma B_i = \Sigma A_i$. Thus each component of R is the sum of a sequence of compact connected sets, each being contained in the interior of the next set of the sequence.

LEMMA 11. 1. *If $\{\Gamma_a{}^1\}$ is a compact 1-cocycle of the locally connected locally compact paracompact normal space R there is a mapping f of R into the circumference S^1 such that $\{\chi^*{}_a f^* \Delta^1\} \sim \{\Gamma_a{}^1\}$ in R.*

Proof. Let $V_\beta \in \mathfrak{V}$ be a component of R and let $\{B_{\beta i}\}$ be a sequence of

compact connected sets such that $B_{\beta i} \subset B_{\beta,i+1}$ and $V_\beta = \Sigma_i B_{\beta i}$. Then, by Theorem 8. 1, there is a mapping $f_{\beta 1}$ of $B_{\beta 1}$ into S^1 such that $f^*_{\beta 1}\Delta^1 \sim \Gamma^1_{\beta 1}$ in $B_{\beta 1}$, where $\Gamma^1_{\beta 1}$ is the cocycle of $\{\Gamma_\alpha{}^1\}$ in the compact set $B_{\beta 1}$. Assume we have a mapping $f_{\beta i}$ of $B_{\beta i}$ into S^1 such that $f^*_{\beta i}\Delta^1 \sim \Gamma^1_{\beta i}$ in $B_{\beta i}$. There exists a mapping $g_{\beta,i+1}$ of $B_{\beta,i+1}$ into S^1 such that $g^*_{\beta,i+1}\Delta^1 \sim \Gamma^1_{\beta,i+1}$ in $B_{\beta,i+1}$. If ψ_i is the identical mapping of $B_{\alpha i}$ in $B_{\alpha,i+1}$, then $\psi^*_i g^*_{\beta,i+1}\Delta^1 \sim \psi^*_i \Gamma^1_{\beta,i+1} \sim \Gamma^1_{\beta,i} \sim f^*_{\beta i}\Delta^1$ in $B_{\beta i}$. Hence $g_{\beta,i+1}\psi_i = g_{\beta,i+1} \mid B_{\beta i}$ is homotopic to $f_{\beta i}$. Hence, by Theorem 10. 2, $f_{\beta i}$ can be extended to a mapping $f_{\beta,i+1}$ of $B_{\beta,i+1}$ into S^1, with $f_{\beta,i+1}$ homotopic to $g_{\beta,i+1}$. Hence $f^*_{\beta,i+1}\Delta^1 \sim g^*_{\beta,i+1}\Delta^1 \sim \Gamma^1_{\beta,i+1}$ in $B_{\beta,i+1}$. Thus there is a sequence of mappings $f_{\beta i}$ of the sets $B_{\beta i}$ in S^1 such that $f^*_{\beta i}\Delta^1 \sim \Gamma^1_{\beta i}$ in $B_{\beta i}$ and such that $f_{\beta,i+1}$ is an extension of $f_{\beta i}$. Since $B_{\beta i} \subset B_{\beta,i+1}$, the limit mapping f_β of V_β in S^1 is continuous. The mapping f of R is S^1 which coincides with f_β in each V_β is the required mapping.

Let C_α be a set consisting of a finite sum of sets $B_{\beta i}$, $B_{\gamma j}$, etc. each from a different component of R. Then there is a cocycle $\Gamma'_\alpha{}^1$ of C_α uniquely determined up to cohomology such that for each of the sets $B_{\beta i}$, if $\psi_\alpha{}^{\beta i}$ is the identical mapping of $B_{\beta i}$ in C_α, $\psi^*_\alpha{}^{\beta i}\Gamma'_\alpha{}^1 \sim \Gamma'_{\beta i}$ in $B_{\beta i}$. The cocycle $\Gamma'_\alpha{}^1$ can be found by formally adding the cocycles $\Gamma^1_{\beta i}$ of the sets $B_{\beta i}$ in the nerve of each covering which separates these sets. Since $\Gamma'_\alpha{}^1$ is unique, $\Gamma'_\alpha{}^1 \sim \Gamma_\alpha{}^1$ in C_α and $\Gamma'_\alpha{}^1 \sim \chi^*_\alpha f^*\Delta^1$ in C_α and hence $\Gamma_\alpha{}^1 \sim \chi^*_\alpha f^*\Delta^1$ in C_α. Now the interiors of the sets $B_{\beta i}$ form a covering of R by open sets and hence any compact set is contained in a finite number of these. Hence any compact set of R is contained in a set C_α consisting of a finite sum of sets $B_{\beta i}$. Hence the sets C_α form a cofinal family of compact sets and thus $\Gamma_\alpha{}^1 \sim \chi^*_\alpha f^*\Delta^1$ in each of the compact sets C_α of a cofinal family. Hence $\{\Gamma_\alpha{}^1\} \sim \{\chi^*_\alpha f^*\Delta^1\}$. This completes the proof of the lemma.

LEMMA 11. 2. *If f is a mapping of the locally connected locally compact paracompact normal space R into the circumference S^1 such that $\{\chi^*_\alpha f^*\Delta^1\} \sim 0$ in R, then f is inessential.*

Proof. Since $\{\chi^*_\alpha f^*\Delta^1\} \sim 0$ in R, $\chi^*_\alpha f^*\Delta^1 \sim 0$ in C_α, and hence for each compact set C_α of R the partial mapping $f \mid C_\alpha$ is inessential. In particular the partial mappings $f \mid B_{\beta i}$ of the sets $B_{\beta i}$ into S^1 are inessential. Hence there is a real function $\phi_{\beta i}$ on $B_{\beta i}$ such that for $x \in B_{\beta i}$, $f(x) = \exp(2\pi i \phi_{\beta i}(x))$. If a point x_β is chosen in $B_{\beta 1}$, $\phi_{\beta i}$ can be chosen so that $0 \leq \phi_{\beta i}(x_\beta) < 1$. Then, since $B_{\beta i}$ is connected, $\phi_{\beta i}$ is uniquely determined.[22] Hence $\phi_{\beta,i+1}$ is an extension of $\phi_{\beta i}$. Hence there is a limit function ϕ_β defined and continuous on V_β such that for $x \in V_\beta$, $f(x) = \exp(2\pi i \phi_\beta(x))$. The function ϕ which

[22] See [13], p. 64

coincides with ϕ_β on V_β for each component V_β is then continuous and $f(x) = \exp(2\pi i\phi(x))$ for $x \in R$. Hence the mapping f of R in S^1 is inessential.

THEOREM 11.3. *If a paracompact normal space R is locally compact and locally connected, the group of homotopy classes of mappings of R into the circumference S^1 is isomorphic with the* 1*-cohomology group of R based on compact cocycles.*

Proof. To each mapping f of R into S^1 corresponds a compact cocycle $\{\chi^*_a f^* \Delta^1\}$. If f_0 and f_1 are homotopic mappings of R into S^1, $f_0 \mid C_a$ and $f_1 \mid C_a$ are homotopic mappings of C_a in S^1, and hence $\chi^*_a f^*_0 \Delta^1 \sim \chi^*_a f^*_1 \Delta^1$ in C_a, and hence $\{\chi^*_a f^*_0 \Delta^1\} \sim \{\chi^*_a f^*_1 \Delta^1\}$ in R. If f is the difference of the two mappings f_0 and f_1 of R into S^1, $f \mid C_a = f_0 \mid C_a - f_1 \mid C_a$, and hence $\chi^*_a f^* \Delta^1 \sim \chi^*_a f^*_0 \Delta^1 - \chi^*_a f^*_1 \Delta^1$ in C_a, and hence $\{\chi^*_a f^* \Delta^1\} \sim \{\chi^*_a f^*_0 \Delta^1\} - \{\chi^*_a f^*_1 \Delta^1\}$ in R. Thus there is a homomorphism of the group of homotopy classes into the cohomology group. But, by Lemma 11.1, there is a mapping for every cocycle. Hence the homomorphism is onto the cohomology group. By Lemma 11.2, only the class of inessential mappings corresponds to the zero element of the cohomology group. Hence the homomorphism is an isomorphism. This completes the proof.

It is easily seen by counterexamples that the requirements that R be locally connected and locally compact are essential for Theorem 11.3. And even if these conditions are satisfied the n-cohomology group of R need not be isomorphic with the n-cohomology group based on compact cocycles for $n > 1$. This may be shown by the following example.

Let R be the space defined as follows: Let the end circles of a sequence of cylindrical surfaces be distinguished as right and left. Let each point of the right circle of the cylinder R_i be identified with its diametrically opposite point and let the circle so obtained be identified with the left circle of R_{i+1}.

The space R can be subdivided to form an infinite complex K. Such subdivisions are obtained by subdividing each of the cylinders R_i so that the subdivisions agree on the common circles. If K_a is any such complex the covering $\mathfrak{U}_a$ of R by the stars of the vertices of K_a has K_a as nerve. Such coverings $\mathfrak{U}_a$ form a cofinal family of coverings of R. Let σ^2_{ai} be a designated 2-simplex of K_a contained in R_i and let σ^1_{ai} be a designated 1-simplex of K_a in the left circle of R_i. Then each 2-cocycle of K_a is cohomologous to a cocycle $c_a{}^i \sigma^2_{ai}$, and it is cohomologous to zero if and only if there exist numbers $a_a{}^i$ such that $c_a{}^i \sigma^2_{ai} = a_a{}^i C^2_{ai}$ where, if $r > 1$, $C^2_{ai} = -2\sigma^2_{a,i-1} + \sigma^2_{ai}$, and $C^2_{a1} = \sigma^2_{a1}$. We assume that K_a is oriented so that C^2_{a1} is cohomologous to the coboundary of σ^1_{ai}.

Let Γ_a^2 be the 2-cocycle $c_a{}^i\sigma^2{}_{ai}$ of K_a, where $c_a{}^i = 1$ or 0 according as i is odd or even. Then $\Gamma^2 = \{\Gamma_a^2\}$ is a cocycle of the cofinal family of coverings. Assume if possible that $\Gamma^2 \sim 0$ in R. Then $\Gamma_a^2 \sim 0$ in K_a for some covering $\mathfrak{U}_a$ of the complete family. Then there are numbers $a_a{}^i$ such that

$$\Gamma_a^2 = a_a{}^i C^2{}_{ai} = a_a{}^1\sigma^2{}_{a1} + \sum_{i=2}^{\infty}(-2a_a{}^i\sigma^2{}_{a,i-1} + a_a{}^i\sigma^2{}_{ai}) = \sum_{i=1}^{\infty}(-2a_a{}^{i+1} + a_a{}^i)\sigma^2{}_{ai}.$$

Hence

$$c_a{}^i = -2a_a{}^{i+1} + a_a{}^i, \text{ i. e., } -2a_a{}^{2n+1} + a_a{}^{2n} = 0, \quad -2a_a{}^{2n+2} + a_a{}^{2n+1} = 1.$$

Therefore

$$2^{2n+1}a_a{}^{2n+2} = 2^{2n-1}a_a{}^{2n} - 2^{2n} = 2a_a{}^2 - \sum_{p=0}^{n-1} 2^{2(n-p)}$$

$$= a_a{}^1 - \sum_{p=0}^{n} 2^{2(n-p)} = a_a{}^1 - \tfrac{1}{3}(2^{2n+2} - 1).$$

Therefore

$$a_a{}^1 \equiv \tfrac{1}{3}(2^{2n+2} - 1) \bmod 2^{2n+1}.$$

Therefore

$$a_a{}^1 \equiv \tfrac{1}{3}(2^{2n+2} - 1) - 2^{2n+1} = -\tfrac{1}{3}(2^{2n+1} + 1) \bmod 2^{2n+1}.$$

Therefore

$$|a_a{}^1| \geq \tfrac{1}{3}(2^{2n+1} - 1) \geq 2^{2n-1} \text{ for every } n.$$

But this is impossible. Hence Γ^2 is not ~ 0 in R.

Therefore not all 2-cocycles of R are cohomologous to zero in R and hence R has essential mappings on S^1. However in a compact subset of R consisting of a finite number of the elementary cylinders (with indentifications) all 2-cocycles are ~ 0. But these compact subsets form a cofinal family of compact subsets of R. Hence every compact 2-cocycle of R is ~ 0. Therefore the 2-cohomology group of R based on infinite coverings is not isomorphic with the one based on compact cocycles.

TUFTS COLLEGE,
MEDFORD, MASS.

BIBLIOGRAPHY

1. P. Alexandroff, "Untersuchungen über Gestalt und Lage abgeschlossener Mengen beliebiger Dimension," *Annals of Mathematics* (2), vol. 30 (1929), pp. 101-187.
2. P. Alexandroff, "Dimensionstheorie," *Mathematische Annalen*, vol. 106 (1932), pp. 161-238.
3. P. Alexandroff and H. Hopf, *Topologie*, Berlin (1935)

4. K. Borsuk, "Über eine Klasse von lokal zusammenhängenden Räumen," *Fundamenta Mathematicae*, vol. 19 (1932), pp. 220-242.
5. K. Borsuk, "Sur les prolongements des transformations continues," *Fundamenta Mathematicae*, vol. 28 (1937), pp. 99-110.
6. K. Borsuk and S. Eilenberg, "Über stetige Abbildungen der Teilmengen euklidischer Räume auf die Kreislinie," *Fundamenta Mathematicae*, vol. 26 (1936), pp. 207-223.
7. N. Bruschlinsky, "Stetige Abbildungen und Bettische Gruppen der Dimensionszahlen 1 und 3," *Mathematische Annalen*, vol. 109 (1934), pp. 525-537.
8. E. Čech, "Théorie générale de l'homologie dans un espace quelconque," *Fundamenta Mathematicae*, vol. 19 (1932), pp. 149-183.
9. E. Čech, "Sur la connexité locale d'ordre supérieur," *Compositio Mathematica*, vol. 2 (1935), pp. 1-25.
10. E. Čech, "On bicompact spaces," *Annals of Mathematics* (2), vol. 38 (1937), pp. 823-844.
11. J. Dieudonné, "Une généralisation des espaces compacts," *Journal de Mathématiques pures et appliquées*, vol. 23 (1944), pp. 65-76.
12. C. H. Dowker, "Hopf's theorem for non-compact spaces," *Proceedings of the National Academy of Sciences*, vol. 23 (1937), pp. 293-294.
13. S. Eilenberg, "Transformations continues en circonférence et la topologie du plan," *Fundamenta Mathematicae*, vol. 26 (1936), pp. 61-112.
14. S. Eilenberg, "Bemerkungen zur Pontrjaginschen Verallgemeinerung des Alexanderschen Dualitätssatzes," *Fundamenta Mathematicae*, vol. 26 (1936), pp. 224-228.
15. S. Eilenberg, "Extension and classification of continuous mappings," *Lectures in Topology*, Ann Arbor (1941), pp. 57-99.
16. H. Freudenthal, "Die Hopfsche Gruppe," *Compositio Mathematica*, vol. 2 (1935), pp. 134-162.
17. H. Hopf, "Die Klassen der Abbildungen der n-dimensionalen Polyeder auf die n-dimensionale Sphäre," *Commentarii Mathematici Helvetici*, vol. 5 (1933), pp. 39-54
18. W. Hurewicz, "Über Abbildungen von endlichdimensionalen Räumen auf Teilmengen Cartesischen Räume," *Sitzungsberichte der Preussischen Akademie der Wissenschaften* (1933), pp. 754-768.
19. W. Hurewicz, "Sur la dimension des produits Cartésiens," *Annals of Mathematics* (2), vol. 36 (1935), pp. 194-197.
20. W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton (1941).
21. S. Kaplan, "Homology properties of arbitrary subsets of Euclidean spaces," to appear in the *Transactions of the American Mathematical Society*.
22. C. Kuratowski, "Sur un théorème fondamental concernant le nerf d'un système d'ensembles," *Fundamenta Mathematicae*, vol. 20 (1933), pp. 191-196.
23. S. Lefschetz, "Abstract complexes," *Lectures in Topology*, Ann Arbor (1941), pp. 1-28.
24. S. Lefschetz, *Algebraic Topology*, New York (1942).
25. S. Lefschetz, *Topics in Topology*, Princeton (1942).
26. A. Weil, "Sur les espaces à structure uniforme et sur la topologie générale" (*Actualités Scientifiques et Industrielles*, No. 551), Paris (1938).
27. H. Whitney, "On matrices of integers and combinatorial topology," *Duke Mathematical Journal*, vol. 3 (1937), pp. 35-45.
28. H. Whitney, "The maps of an n-complex into an n-sphere," *Duke Mathematical Journal*, vol. 3 (1937), pp. 51-55.

ON THE ZETA-FUNCTIONS OF ALGEBRAIC NUMBER FIELDS.*

By RICHARD BRAUER.

1. It was proved by E. Artin[1] that if k is an algebraic number field (of finite degree) and K a normal extension field with the icosahedral group as the Galois group with regard to k, then the zeta-function $\zeta(s,k)$ of k divides the zeta-function $\zeta(s,K)$, in the sense that the quotient $\zeta(s,K)/\zeta(s,k)$ is an integral function of the complex variable s. Using Artin's method, we shall show in this note that *the zeta-function $\zeta(s,k)$ of an algebraic number field k divides the zeta-function $\zeta(s,K)$ of every normal extension field K of k.*[1a] The proof (**3**) is based on a group-theoretical lemma which is established in **2**. Our result will enable us to prove in **4** the following theorem conjectured by C. L. Siegel.[2] *Consider all algebraic number fields of a fixed degree n. Let d be the discriminant of k, let h be the number of classes of ideals in k, and let R be the regulator of k. Then*

$$(*) \qquad \log(hR) \sim \log\sqrt{|d|}, \qquad (\text{for } |d| \to \infty).$$

This had been proved by Siegel for quadratic fields k. In the case of imaginary quadratic fields k, the result gives a refinement of the famous theorem of Heilbronn that h tends to infinity with $|d|$. Siegel also mentions that his method makes it possible to establish (*) for all fields k of fixed degree n which are *soluble* with regard to a fixed subfield. Using the above results on the zeta-function, and some results of Landau, we can apply Siegel's original method to obtain (*) in its full generality.

2. Let G be a group of finite order g. The irreducible characters of G will be denoted by $\chi_0, \chi_1, \cdots, \chi_{n-1}$ where in particular χ_0 is the 1-character, $\chi_0(\sigma) = 1$ for all σ in G. If f_ν is the degree of χ_ν, the character θ of the regular representation of G is given by

* Received January 9, 1947; Presented to the American Mathematical Society, April 25, 1947.

[1] E. Artin, *Mathematische Annalen*, vol. 89 (1923), pp. 147-156; *Abhandlungen aus dem Mathematischen Seminar Hamburg*, vol. 3 (1924), pp. 89-108. Cf. also E. Artin, *Abhandlungen aus dem Mathematischen Seminar Hamburg*, vol. 8 (1931), pp. 292-306.

[1a] After I had submitted the present paper, H. W. Brinkmann kindly drew my attention to the fact that this Theorem 1 had already been obtained by Hideo Aramata, *Proceedings of the Imperial Academy of Japan*, vol. 9 (1933), pp. 31-34. I publish my proof in the following since it seems to be somewhat simpler.

[2] C. L. Siegel, *Acta Arithmetica*, vol. 1 (1935), pp. 83-86.

$$\theta = \sum_{\nu=0}^{n-1} f_\nu \chi_\nu. \tag{1}$$

LEMMA. *The character*

$$\theta - \chi_0 = \sum_{\nu=1}^{n-1} f_\nu \chi_\nu$$

of G can be expressed as a linear combination $\Sigma c_\rho \omega_\rho^*$ of characters ω_ρ^* of G which are induced by irreducible characters ω_ρ different from the 1-character of cyclic subgroups, such that the coefficients c_ρ are positive rational numbers with the denominator g.

Proof. Let A be a cyclic subgroup of order $a > 1$ of G. Let $\psi_0(\alpha) = 1$, $\psi_1(\alpha), \cdots, \psi_{a-1}(\alpha)$ be the characters of A where α denotes a variable element of A. Then $\chi_\nu(\alpha)$ can be written as a linear combination

$$\chi_\nu(\alpha) = \sum_{\mu=0}^{a-1} h_{\nu\mu} \psi_\mu(\alpha) \tag{2}$$

with coefficients $h_{\nu\mu}$ which are non-negative rational integers. The orthogonality relations for group characters show that here

$$h_{\nu\mu} = (1/a) \sum_{\alpha} \overline{\chi_\nu(\alpha)} \psi_\mu(\alpha), \tag{3}$$

α ranging over all elements of A.

The character ψ_μ of A induces a character ψ_μ^* of G. According to a theorem of Frobenius,[3] we have

$$\psi_\mu^*(\sigma) = \sum_{\nu=0}^{n-1} h_{\nu\mu} \chi_\nu(\sigma) \qquad (\sigma \text{ in } G) \tag{4}$$

where the coefficients $h_{\nu\mu}$ in (4) are the same numbers as in (2).

Thus

$$\psi_\mu^*(\sigma) = (1/a) \sum_{\nu=0}^{n-1} \chi_\nu(\sigma) \sum_{\alpha} \overline{\chi_\nu(\alpha)} \psi_\mu(\alpha). \tag{5}$$

Determine now numbers $u_0, u_1, \cdots, u_{a-1}$ such that

$$\sum_{\mu=0}^{a-1} u_\mu \psi_\mu(\alpha) = \begin{cases} a\phi(a), \text{ if } \alpha = 1.[4] \\ 0, \text{ if } \alpha \neq 1, \text{ but does not generate } A. \\ -a, \text{ if } \alpha \text{ generates } A. \end{cases} \tag{6}$$

The orthogonality relations for the group characters of A make is possible to write down the u_μ explicitly. We have

---

[3] G. Frobenius, *Sitzungsberichte der Berliner Akademie* (1898), pp. 501-515. Compare also A. Speiser, *Theorie der Gruppen von endlicher Ordnung*, 3rd ed. (1937), Berlin, § 64.

[4] Here, $\phi(a)$ denotes the Euler function.

$$u_\mu = (1/a)\,(\phi(a)\overline{a\psi_\mu(1)} + \sum_\beta (-a)\overline{\psi_\mu(\beta)})$$

where β ranges over the primitive elements of A. Hence

$$u_\mu = \phi(a) - \sum_\beta \overline{\psi_\mu(\beta)}. \tag{7}$$

This shows that u_μ is a rational number and hence a rational integer. Further, since $|\psi_\mu(\beta)| = 1$ and since $\psi_\mu(\beta) \neq 1$ except for the 1-character ψ_0, it follows from (7) that

$$u_0 = 0,\ u_\mu > 0 \text{ for } \mu > 0. \tag{8}$$

Form now the expression

$$S(A) = \sum_\mu u_\mu \psi_\mu^*(\sigma). \tag{9}$$

Then, by (5)

$$S(A) = (1/a) \sum_\nu \chi_\nu(\sigma) \sum_\alpha \overline{\chi_\nu(\alpha)} \sum_\mu u_\mu \psi_\mu(\alpha)$$

and (6) yields

$$S(A) = \sum_\nu \chi_\nu(\sigma)\,(\overline{\chi_\nu(1)}\phi(a) - \sum_\beta \overline{\chi_\nu(\beta)})$$

where β again ranges over the primitive elements of A. This shows that the function $S(A)$ of a variable element σ of G is a linear combination of the characters χ_ν. The value $\overline{\chi_\nu(1)}$ is the degree f_ν of χ_ν. In particular, the 1-character χ_0 appears in $S(A)$ with the coefficient

$$\phi(a) - \sum_\beta 1 = \phi(a) - \phi(a) = 0.$$

It is sufficient to let ν range over the values $1, 2, \cdots, n-1$:

$$S(A) = \sum_{\nu=1}^{n-1} \chi_\nu(\sigma)\,(f_\nu\phi(a) - \sum_\beta \overline{\chi_\nu(\beta)}). \tag{10}$$

Add now the formulas (10) for *all* cyclic subgroups A of an order > 1. Then β will range over all elements $\neq 1$ of G, and since χ_ν is not the 1-character χ_0, the orthogonality relations for group characters yield

$$-\sum_A \sum_\beta \overline{\chi_\nu(\beta)} = \overline{\chi_\nu(1)} = f_\nu.$$

Further

$$\sum_A \phi(a) = g - 1.$$

Hence (10) gives

$$\sum S(A) = \sum_{\nu=1}^{n-1} \chi_\nu(\sigma)\,(f_\nu(g-1) + f_\nu) = g \sum_{\nu=1}^{n-1} f_\nu \chi_\nu(\sigma).$$

Thus

$$(11) \qquad (1/g)\sum S(A) = \sum_{\nu=1}^{n-1} f_\nu \chi_\nu(\sigma) = \theta - \chi_0.$$

It is shown by (9) that the left side of (11) is a linear combination of characters ψ_μ^* induced by irreducible characters ψ_μ of cyclic subgroups. From (8), it follows that no character ψ_0^* induced by a 1-character appears, and that the other ψ_μ^* appear with positive coefficients u_μ/g. These coefficients are rational numbers with the denominator g. This establishes the lemma.

3. We now state

THEOREM 1. *Let k be an algebraic number field and let K be a normal field over k, both fields being of finite degrees. Then the quotient of the zeta-functions*

$$\zeta(s, K)/\zeta(s, k)$$

is an integral function of the complex variable s.

Proof. The zeta-functions can be expressed as L-series in the sense of Artin [5]

$$\zeta(s, K) = L(s; \theta, K/k)\,; \zeta(s, k) = L(s; \chi_0, K/k)$$

where θ now denotes the character of the regular representation of the Galois group G of K with regard to k and where χ_0 is the 1-character of G. Hence

$$\zeta(s, K)/\zeta(s, k) = L(s; \theta - \chi_0, K/k)$$

and Lemma 1 yields

$$(12) \qquad \zeta(s, K)/\zeta(s, k) = \Pi L(s; \omega_\rho^*, K/k)^{c_\rho}.$$

According to a fundamental result of Artin,[6] the L-series $L(s; \omega_\rho^*, K/k)$ is identical with an abelian L-series $L(s, \omega_\rho) = L(s; \omega_\rho, U/V)$ where U is a cyclic field over the field V; $k \subseteq V \subset U \subseteq K$. Here, ω_ρ is a linear character, not the 1-character, and consequently [7] $L(s; \omega_\rho^*, K/k)$ is an integral function. Now (12) together with Lemma 1 shows that the g-th power of $\zeta(s, K)/\zeta(s, k)$ is an integral function. Since $\zeta(s, K)/\zeta(s, k)$ is meromorphic, it must be integral. This proves Theorem 1. At the same time we obtain:

---

[5] See § 4 of Artin's paper in the *Abhandlungen aus dem Mathematischen Seminar Hamburg*, vol. 3.

[6] See § 5 of Artin's paper referred to in [5]. The general law of reciprocity was proved by Artin in *Abhandlungen aus dem Mathematischen Seminar Hamburg*, vol. 5 (1927), pp. 353-363.

[7] This had first been shown by E. Hecke. See E. Landau, *Mathematische Zeitschrift*, vol. 2 (1918), pp. 52-154, Theorem LXIII.

COROLLARY. *If the normal field K has the degree g over k, then the g-th power of $\zeta(s, K)/\zeta(s, k)$ can be written as a product of abelian L-series which belong to linear characters different from the* 1*-character.*

It has been surmised that Theorem 1 remains valid, if K is an arbitrary (not normal) extension field of k of finite degree. This would follow from the still stronger conjecture of Artin that the L-functions belonging to irreducible characters different from the 1-character are integral functions. Our present method does not offer a possibility to attack these problems.[8]

4. We now wish to prove the following theorem formulated by Siegel.

THEOREM 2. *Let d denote the discriminant of the algebraic number field k of degree n over the field P of rational numbers. Let h be the number of classes of ideals in k and let R be the regulator. Then, for the fields k of fixed degree n*

$$\log(hR) \sim \log \sqrt{|d|} \qquad (\text{for } |d| \to \infty).$$

Proof. (a) We state without proof a number of known results which will be used in what follows.

LEMMA 1 (Siegel[9]). *Let W be an algebraic number field of degree q over the field of rational numbers with the discriminant T. Denote by $r(W)$ the residue of the zeta-function $\zeta(s, W)$ for $s = 1$. If $\zeta(s, W) \leqq 0$ for a real value of s with $0 < s < 1$, then*

$$r(W) > s(1-s)2^{-q}e^{-2q\pi}\,|T|^{(s-1)/2}. \tag{13}$$

Siegel's Lemma 4 must be replaced by

LEMMA 2 (Landau[10]). *Let W, q, T have the same significance as in Lemma* 1. *Let χ be a linear character, not the* 1*-character, modulo the ideal $\mathfrak{f}$ in W. The corresponding (abelian) L-series $L(s, \chi)$ satisfies for $s = 1$ the inequality*

---

[8] In the case of the icosahedral group where the characters of G are known explicitly, Artin could show that the zeta-function of the ground field divides the zeta-functions of some of the fields Ω lying between k and K, but not all these Ω can be treated in this manner. Actually, it is not known to-day whether $\zeta(s, k)$ divides $\zeta(s, \Omega)$ in these other cases.

[9] Siegel, *loc. cit.*, Lemma 2. The proof of the lemma, as is that of the following two lemmas, is based on Hecke's integrals for the zeta-functions and the abelian L-series.

[10] E. Landau, *Mathematische Zeitschrift*, vol. 4 (1919), pp. 152-162; theorem 5. While the corresponding lemma in Siegel's case can be proved quite directly in an elementary manner, it seems impossible at the present time to prove Lemma 3 avoiding Hecke's theory of analytic continuation of the abelian L-series.

$$|L(1,\chi)| \leqq \lambda_q \log^q |T \cdot N\mathfrak{f}| \tag{14}$$

where λ_q depends on q alone and where $N\mathfrak{f}$ denotes the norm of $\mathfrak{f}$.

LEMMA 3 (Landau [11]). *Let k be an algebraic number field of degree n over the field of rational numbers. If h denotes the number of classes of ideals in k, if d is the discriminant and R the regulator of k, then*

$$hR \leqq \lambda'_n \sqrt{|d|} \log^{n-1} |d| \tag{15}$$

where λ'_n is a constant depending on n alone.

(b) We wish to show that for all algebraic number fields k of degree n the relation

$$\log r(k) = o(\log |d|) \tag{16}$$

holds where $r(k)$ is the residue of $\zeta(s,k)$ for $s=1$ and where d is the discriminant of k. As is well known,

$$r(k) = chR/\sqrt{|d|} \tag{17}$$

where c lies between positive bounds depending on n alone. Hence (15) shows that

$$\overline{\lim} \log r(k)/\log |d| \leqq 0. \tag{18}$$

We now consider the case where k is a normal field of degree n over the field of rational numbers and apply Siegel's method. If (16) were false for normal fields k, there would exist a positive number $\epsilon < 1$ such that for infinitely many normal fields k of degree n the following inequality holds

$$r(k) < |d|^{-\epsilon}. \tag{19}$$

Set

$$\eta = 2\epsilon/n \tag{20}$$

and choose k such that (19) holds and that $|d|$ satisfies the inequality

$$|d|^{-\epsilon} < (1-\eta)\eta 2^{-n} e^{-2n\pi} |d|^{-\eta/2}. \tag{21}$$

By Lemma 1 and (19) and (21), $\zeta(1-\eta,k) > 0$. Since $\zeta(s,k) \to -\infty$ for real s approaching 1 from the left, it follows that there exists a real σ with

[11] E. Landau, *Göttingen Nachrichten* (1918), pp. 478-488, proof of Lemma 1.

$$1-\eta<\sigma<1 \tag{22}$$

such that

$$\zeta(\sigma, k)=0. \tag{23}$$

From now on, the field k will be considered as fixed. Let K be a second field, normal of degree n over the field of rational numbers, and let $\Omega = Kk$ be the compositum of K and k. Then the degree m of Ω divides n^2. If K has the discriminant D, the discriminant Δ of Ω will divide $d^n D^n$.

It follows from the Corollary in **3** that we may set

$$\zeta(s, \Omega)=\zeta(s, k) M_0(s), \tag{24}$$

$$\zeta(s, \Omega)=\zeta(s, K) M(s). \tag{25}$$

Here, the n^2-th powers of $M_0(s)$ and $M(s)$ are products of L-series $L(s; \chi, U/V)$ where U is a cyclic extension field of the field V, $V \subset U \subseteq \Omega$, and where χ is a linear character, different from the 1-character. The number of factors lies below a fixed bound depending on n. Then $L(s; \chi, U/V)$ is identical with an L-series $L(s, \chi)$ of the field V where χ now is a character modulo a suitable ideal $\mathfrak{f}$, the conductor, and χ is not the 1-character. It follows from the conductor-discriminant formula of class field theory [12] that $N\mathfrak{f}$ divides the discriminant of U which itself divides the discriminant Δ of Ω and hence $d^n D^n$. The discriminant of V also will divide $d^n D^n$. Take $|D| \geqq |d|$. Lemma 2 implies that

$$|L(1, \chi, U/V)| \leqq c_1 \log^{c_2} |D|$$

where c_1, c_2 (and later $c_3, c_4, \cdots$) are positive constants depending on n alone. Then

$$|M(1)| \leqq c_3 \log^{c_4} |D|. \tag{26}$$

It follows from (23) and (24) that $\zeta(\sigma, \Omega) = 0$. Hence, by Lemma 1,

$$r(\Omega) > \sigma(1-\sigma) 2^{-m} e^{-2m\pi} |\Delta|^{(\sigma-1)/2}$$

where m is the degree of Ω. Since $|\Delta| \leqq |d|^n |D|^n$, this gives

$$r(\Omega) \geqq a |D|^{-n(1-\sigma)/2}, \tag{27}$$

a denoting a positive constant which depends on n, d, σ, but not on K. It follows from (25) that

[12] See, for instance, H. Hasse, *Klassenkörpertheorie*, mimeographed notes, Marburg 1932/33, p. 118 and § 23. Actually, we use only a weaker result. This, as is well known, can be obtained analytically by comparing Hecke's functional equations for $\zeta(s, U)$ and for the L-series belonging to the cyclic field U over V.

$$r(\Omega) = r(K)M(1).$$

Combining this with (26) and (27), we obtain

$$r(K) = r(\Omega)/M(1) \geqq ac_3^{-1} \mid D \mid^{-n(1-\sigma)/2} \log^{-c_4} \mid D \mid .$$

Since $1 - \sigma < \eta = 2\epsilon/n$, it follows that

$$r(K) \geqq \mid D \mid^{-\epsilon}$$

for all normal fields K of degree n with sufficiently large $\mid D \mid$. This contradicts the assumption that the inequality (19) holds for infinitely many normal fields of degree n, and this shows that (16) holds for normal fields k.

(c) Consider now an arbitrary number field k of degree n with the discriminant d. Again, we wish to establish the relation (16). Let Ω denote now the Galois field belonging to k. Then Ω is normal over the field of rational numbers. Its degree lies below a fixed constant c_5 depending on n only. If Δ is the discriminant of Ω, then $\mid \Delta \mid \leqq \mid d \mid^{\nu}$ where ν depends only on n. On account of the Corollary in **3**, we may set

$$\zeta(s, \Omega) = \zeta(s, k)M^*(s) \tag{28}$$

where $M^*(s)$ has properties analogous to those of $M(s)$ in (25). Applying the argument leading to (26) here, we obtain now

$$\mid M^*(1) \mid \leqq c_6 \log^{c_7} \mid d \mid .$$

Now (28) yields

$$r(k) = r(\Omega)/M^*(1) \geqq r(\Omega)c_6^{-1} \log^{-c_7} \mid d \mid . \tag{29}$$

Since (16) has been established for normal fields of a fixed degree, we have $r(\Omega) \geqq \mid \Delta \mid^{-\delta}$ for every fixed $\delta > 0$ provided that $\mid \Delta \mid$ is sufficiently large. But $\mid d \mid \leqq \mid \Delta \mid \leqq \mid d \mid^{\nu}$ and hence

$$r(\Omega) \geqq \mid d \mid^{-\delta\nu} \tag{30}$$

provided that $\mid d \mid$ is sufficiently large. Combination of (29) and (30) gives

$$\underline{\lim} \log r(k)/\log \mid d \mid \geqq 0.$$

This, together with (18), shows that (16) holds for all fields k of degree n.

Now (17) implies that $\log (hR) \sim \log \sqrt{\mid d \mid}$ as was stated in Theorem 2.

UNIVERSITY OF TORONTO.

ASYMPTOTIC INTEGRATIONS OF THE ADIABATIC OSCILLATOR.*

By AUREL WINTNER.

---

1. The following considerations deal with the behavior, as $t \to \infty$, of the linear, non-conservative system defined by the Hamiltonian function

$$H(x, x'; t) = \tfrac{1}{2}\{x'^2 + (\omega^2 + \phi(t))x^2\} \tag{1}$$

of a single degree of freedom, that is, by the Lagrangian equation

$$x'' + (\omega^2 + \phi(t))x = 0, \tag{2}$$

where ω denotes a positive constant ("undisturbed frequency"), and the coefficient function $\phi(t)$, a given perturbation of the squared frequency, is "small" or "slow" as $t \to \infty$. It will always be assumed that the coefficient function, $\phi(t)$, of (2) is defined for large positive t, and that $\phi(t)$ is continuous and real-valued (these assumptions will not always be repeated in the wording of the theorems).

Actually, it will be convenient (and, in the proof of (i), almost necessary) to consider complex-valued coefficient functions also. Such coefficient functions will not be denoted by $\phi(t)$ but by other symbols. It will be understood that these coefficient functions, too, are defined and continuous for large positive t.

The "smallness" of $\phi(t)$, as $t \to \infty$, can be specified in various senses. The problem is precisely which of these specifications leads to an assigned restriction of the asymptotic behavior of the general solution of (2). The most natural of these specifications seems to be

$$\phi(t) \to 0 \text{ as } t \to \infty, \tag{3}$$

but this turns out to be quite useless, since it proves to be compatible with almost any asymptotic behavior of the solutions $x(t)$ (except, of course, with a behavior excluded by Sturm's comparison theorem on nodes); cf. [4]. Less primitive specifications result if $\phi(t)$ is required to be of class $(L) = (L^1)$

---

* Received November 7, 1946.

or of class (L^2). Neither of these specifications implies the other and both of them together do not imply (3). It is understood that, since $\phi(t)$ is continuous on the half-line, it is said to be of class (L^p) if

$$(3_p) \qquad \int^{\infty} |\phi(t)|^p \, dt < \infty \qquad (p > 0).$$

Still another condition is

$$(3_0) \qquad \int^{\infty} |d\phi(t)| < \infty$$

(to be required for a sufficiently large lower limit of integration). Since (3_0), in contrast to (3_p), does not imply that $\phi(t)$ is "small" for large t, the restriction (3) will always be assumed, if (3_p), where $p > 0$, is replaced by (3_0). In this regard, it is sufficient to observe that (3_0) implies the existence of a finite limit $\phi(\infty)$, which can have an arbitrary value, and that (3) means the vanishing of this limit.

It is convenient to think of (3_0) and (3) together as representing a limiting case of (3_p). This limiting case of (L^p) will be denoted by (L^0). In other words, $\phi(t)$ will be called of class (L^0) if it is of bounded variation (on *some* half-line) and tends to 0 as $t \to \infty$. For instance, a monotone $\phi(t)$ is of class (L^0) if and only if it satisfies (3). Needless to say, neither of the classes (L^0), (L^p) contains the other (whether $p > 0$ is fixed or variable). All that is true is that $\phi(t)$ must be of class (L^0) if it satisfies (3_0) and is of class (L^p) (for some $p > 0$).

2. The results to be obtained are *explicit rules for the asymptotic integration of* (2), the description of which can best be centered about the following theorem:

(i) *If $\phi(t)$ is of class (L^0) and of class (L^2), then there belongs to every solution $x(t) \not\equiv 0$ of (2) a unique pair of positive integration constants, say a and α $(\leqq 2\pi)$, such that*

$$(4) \qquad x(t) = a \cos\left(\alpha + \omega t + \tfrac{1}{2}\int^{t} \phi(s)\, ds/\omega\right) + \epsilon(t),$$

where the remainder term, $\epsilon(t) = \epsilon_{a\alpha}(t)$, is $o(1)$, as is its derivative; that is,

$$(5) \qquad \epsilon(t) \to 0 \text{ and } \epsilon'(t) \to 0 \text{ as } t \to \infty.$$

If the zeros of the cosine in (4) are disregarded, the first of the relations

(5) states that the first term on the right of (4) is asymptotic to the function on the left of (4), and the second of the relations (5) means that the formal differentiation of this asymptotic formula for $x(t)$ is legitimate.

Since (4) contains two integration constants, a and α, the lower limit of the quadrature occurring in (4) can be thought to be fixed. On the other hand, the fact that this lower limit is arbitrary cannot dispose of the necessity of an α in (4), since the function of t representing the third term beneath the cos can have a range the length of which is less than 2π. But this precaution is surely superfluous unless the coefficient function $\phi(t)$ of (2) is such that

$$\int^t \phi(s)\, ds \tag{6}$$

becomes $O(1)$ as $t \to \infty$.

For the still more restrictive case, in which (6) tends to a finite limit, (i) clearly contains the following corollary:

(i bis) *If $\phi(t)$ is of finite total variation and such as to make both improper integrals*

$$\int^\infty \phi^2(t)\,dt, \qquad \int^\infty \phi(t)\,dt \equiv \lim_{T\to\infty} \int^T \phi(t)\,dt \tag{7}$$

convergent (whereas $\int^\infty |\phi(t)|\,dt = \infty$ is allowed), then there belongs to every solution, $x(t)$, of (2) a unique pair of integration constants, c_1 and c_2, such that

$$x(t) = c_1 \cos \omega t + c_2 \sin \omega t + \epsilon(t), \tag{8}$$

where the remainder term and its derivative satisfy (5).

Actually, the (L^2)-assumption can now be omitted, since it is implied by the remaining assumptions of (i bis). In order to see this, it is sufficient to apply to the first of the integrals (7) an integration by parts (based on the factorization $\phi^2(t)\,dt = \phi(t)\,dF(t)$, where F denotes the indefinite integral of ϕ).

What is essential in (i bis) is the inclusion of the case mentioned parenthetically after (7). In fact, if $\phi(t)$ is of class (L^1), then (8) and (5) hold for all solutions of (2) even if neither of the remaining assumptions of (i bis), requiring that $\phi(t)$ be of class (L^0) and of class (L^2), is satisfied. Cf. the case $f(t) \equiv \omega^2$ of (iii) below.

A simple instance, to which (i) is applicable but (i bis) is not, results on choosing $\phi(t) = t^{-\lambda}$, where $\frac{1}{2} < \lambda \leqq 1$. In fact, $\phi(t)$ then is of class (L^0) and of class (L^2), but (6) is not $O(1)$. According to (4), the general solution of

$$x'' + (1 + t^{-\lambda})x = 0 \tag{9}$$

is of the form

$$x(t) = c_1 \cos(t + \tfrac{1}{2}\log t) + c_2 \sin(t + \tfrac{1}{2}\log t) + \epsilon(t) \tag{10}$$

if $\lambda = 1$, and of the form

$$x(t) = c_1 \cos(t + t^{\mu}/\mu) + c_2 \sin(t + t^{\mu}/\mu) + \epsilon(t), \text{ where } \mu = 1 - \lambda, \tag{11}$$

if $\frac{1}{2} < \lambda \neq 1$; hence, of the form (8) if and only if $\mu < 0$, i. e., $\lambda > 1$.

3. Instead of the limiting case of the range $\frac{1}{2} < \lambda \leqq 1$, covered above, consider the following modification of the case $\lambda = \frac{1}{2}$ of (9):

$$x'' + (1 + 2t^{-\frac{1}{2}} + t^{-1})x = 0; \tag{12}$$

so that, in (2),

$$\phi(t) = 2t^{-\frac{1}{2}} + t^{-1} \tag{13}$$

and $\omega = 1$. Since (13) means that

$$\tfrac{1}{2}\int^t \phi(s)\,ds = 2t^{\frac{1}{2}} + \tfrac{1}{2}\log t, \tag{14}$$

the assumption that (i) is applicable would mean that the general solution of (12) is a superposition of the real and imaginary parts of

$$\exp i(t + 2t^{\frac{1}{2}} + \tfrac{1}{2}\log t) + \epsilon(t). \tag{15}$$

In fact, (14) shows that this is equivalent to the assertion, (4), of (i), the value of ω being 1. But it will be shown that the general solution of (12) is a superposition of the real and imaginary parts of

$$\exp i(t + 2t^{\frac{1}{2}}) + \epsilon(t), \tag{16}$$

rather than of (15). It is understood that $\epsilon(t)$ in (15) and (16) denote two functions which (along with their derivative) satisfy (5). Hence, no superposition ($\not\equiv 0$) resulting from (16) can result from (15). Consequently, the assumption that (i) is applicable to (12) is false.

Since the function (13) is of class (L^0) but fails to be (though just barely) of class (L^2), it follows that the second of the assumptions of (i) cannot be omitted:

(i*) *The assertions of* (i) *are not in general true for a coefficient function,* $\phi(t)$, *which is of class* (L^0) (*but not, of course, of class* (L^2) *as well*).

In order to prove this, it will be convenient to start with the principal term,

$$x(t) = \exp i(t + 2t^{\frac{1}{2}}), \tag{16*}$$

of (16) and to deduce for (16*) a differential equation, say (12*), which differs from (12) only in unessential terms (to be specified in a moment), rather than to start with (12) and then deduce (16), which differs from (16*) only in unessential terms (as specified by (5)).

First, if $x = x(t)$ is defined by (16*), then $x' = i(1 + t^{-\frac{1}{2}})x$, hence

$$x'' = i^2(1 + t^{-\frac{1}{2}})^2 x - \tfrac{1}{2}(t^{-\frac{1}{2}})^3 ix.$$

This can be written in the form

$$x'' + (1 + 2t^{-\frac{1}{2}} + t^{-1} + \tfrac{1}{2}(t^{-\frac{1}{2}})^3 i)x = 0, \tag{12*}$$

which is (2) with $\omega = 1$ and with ϕ^* instead of $\phi(t)$, where

$$\phi^*(t) = \phi(t) + \tfrac{1}{2}(t^{-\frac{1}{2}})^3 i, \tag{13*}$$

if $\phi(t)$ denotes the coefficient function, (13), of (12).

Accordingly, the deviation of the coefficient functions of (12) and (12*), though imaginary, is $O(t^{-\frac{1}{2}})^3$, and therefore of class (L^1). But (iii) below will show that a deviation of class (L^1) in the coefficient functions, whether real-valued or not, does not have any effect on the asymptotic form of the general solution; in the sense that the whole disturbance will consist of an additive correction term which (along with its derivative) satisfies (5). Since, in terms of the superposition of real and imaginary parts, (16*) belongs to (12*), it follows that what belongs to (12) is (16), rather than (15).

This proves (i*). In order to ascertain that both assumptions of (i) are essential, the following dual of (i*) remains to be verified:

(i**) *The assertions of* (i) *are not in general true for a coefficient function,* $\phi(t)$, *which tends, as* $t \to \infty$, *to* 0 *and is of class* (L^2) (*but not, of course, of class* (L^0) *as well*).

In fact, Perron [2] has exhibited a (real-valued) coefficient function $\phi(t)$ and a solution $x = x(t)$ of the differential equation (2), belonging to this $\phi(t)$ which have the following properties: $\phi(t) = O(t^{-1})$ and $x(t) \neq O(1)$, as $t \to \infty$ (for a simple deduction and for more general constructions, cf. [4], pp. 386-387 and pp. 394-396). This example proves (i**), since, on the one hand, $\phi(t) = O(t^{-1})$ implies that $\phi(t)$ is of class (L^2) and satisfies (3) and, on the other hand, $x(t) \neq O(1)$ contradicts the assertion, (4), of (i).

4. What is known under the (L^0)-assumption alone can be summarized as follows:

(ii$_0$) *If $\phi(t)$ is of class (L^0), then, although (4) and (5) need not hold,*

$$x(t) = O(1) \text{ and } x'(t) = O(1) \text{ as } t \to \infty \tag{17}$$

must hold, for every solution, $x = x(t)$, of (2).

The first assertion of (ii$_0$) is precisely (i*). The second claims that (3) and (3$_0$) imply (17) for every solution of (2). This has been proved by A. Kneser [1], pp. 73-80. In fact, only his wording, but not his proof, of (17) assumes more than the pair of conditions (3), (3$_0$). But the proof is made unnecessarily involved by Kneser's use of Sturm's oscillation theorem. If the latter is avoided, more than (17) can be proved:

(ii) *If $\phi(t)$ is of class (L^0) and if*

$$h(t) = H(x(t), x'(t); t) \tag{18}$$

denotes the energy, (1), along any solution, $x = x(t)$, of (1), then

$$\int^{\infty} |\, d\, h(t)| < \infty. \tag{19}$$

COROLLARY. *The energy, (18), of every solution $x = x(t)$ tends, as $t \to \infty$, to a finite limit,*

$$h(\infty). \tag{20}$$

The latter is positive except when

$$x(t) \equiv 0, \text{ hence } h(t) \equiv 0. \tag{20 bis}$$

In order to obtain (17), not even the existence of a finite limit, (20), is needed, since $h(t) = O(1)$ is sufficient to this end. In fact, $h(t) = O(1)$ means, by (18) and (1), that

$$x'^2(t) + (\omega^2 + \phi(t))x^2(t) = O(1),$$

which, in view of (3), is equivalent to (17).

Remark. Let a_n denote the n-th amplitude, and b_n the n-th co-amplitude, of a (real-valued) solution $x = x(t) \not\equiv 0$ (as to terminology, cf. [4], pp. 389-391). Then it is easily seen from (21°) below, from the notations (18) and (1), and from (3), that the assertion, (19), of (ii), implies

$$(21) \qquad \sum_{n=1}^{\infty} | a_{n+1} - a_n | < \infty \text{ and } \sum_{n=1}^{\infty} | b_{n+1} - b_n | < \infty.$$

Kneser's result, according to which there exist finite positive limits

$$(21_0) \qquad \lim_{n\to\infty} a_n \quad \text{and} \quad \lim_{n\to\infty} b_n$$

if $x(t) \not\equiv 0$ ([1], p. 80) is, of course, a corollary of (21). Finally,

$$(21°) \qquad \lim_{n\to\infty} a_{n+1}/a_n = 1 \quad \text{and} \quad \lim_{n\to\infty} b_n/a_n = \omega$$

hold even if $\phi(t)$, instead of satisfying both (3) and (3_0), satisfies (3) only (cf. [4], p. 391).

5. The proof of (ii) can be based on the fact that, in virtue of (18) and (1),

$$(22) \qquad dh(t) = \tfrac{1}{2}x^2(t)\,d\phi(t)$$

is an identity in t along any solution, $x = x(t)$, of (2). This fact is a straight-forward generalization of the energy relation, $h(t) = \text{const.}$, of the conservative case, $\phi(t) \equiv 0$, and is verified in the same way, that is, by direct substitution and partial integration. Caution is necessary only if $\phi(t)$ is just continuous. In the sense of Stieltjes integrations, (22) is true in this case also, it being understood that (22) must then be interpreted as follows:

$$(22 \text{ bis}) \qquad \int_u^v g(t)\,dh(t) = \tfrac{1}{2}\int_u^v g(t)x^2(t)\,d\phi(t)$$

holds for arbitrary u, v, and for every $g(t)$ of bounded variation.

It is clear from (18), (1) and (3) that, from a certain t onward,

$$(23) \qquad x^2(t) + x'^2(t) < h(t),$$

if the unit of length on the t-axis is so chosen that $4\omega^2$ becomes 1. Let the

trivial solution, (20 bis), be excluded. Then $x(t)$ and $x'(t)$ cannot vanish simultaneously and so, by (23),

$$h(t) > 0. \tag{23 bis}$$

But (22) and (23) imply that, corresponding to the interpretation, (22 bis), of (22),

$$|\, dh(t)\,| \leqq h(t)\,|\, d\phi(t)\,|\,; \text{ hence, } |\, d \log h(t)\,| \leqq |\, d\phi(t)\,|\,,$$

by (23 bis). It follows, therefore, from (3_0) that

$$\int^{\infty} |\, d \log h(t)\,| < \infty.$$

This implies that $\log h(t)$ tends to a finite limit as $t \to \infty$. In particular, $\log h(t) = O(1)$, and so both $h(t)$ and $1/h(t)$ are $O(1)$. Hence, the result of the last formula line is equivalent to the assertion, (19), of (ii). Finally, the vanishing of (20) is precluded by $1/h(t) = O(1)$; so that $h(\infty) > 0$, by (23 bis).

5 bis. More interesting than the set of relations (21), (21_0), (21^0) is the following fact (which, as will be shown elsewhere, is needed in an application to the theory of spectra):

(ii bis) *If $\phi(t)$ is of class (L^0), and if $t_1, t_2, \cdots$ denotes the sequence of consecutive zeros of an arbitrary solution $x(t) \not\equiv 0$ of (2), then the mean square of $x(t)$ over a half-wave, that is, the average*

$$\int_{t_n}^{t_{n+1}} x^2(t)\,dt/(t_{n+1} - t_n),$$

(as well as $t_{n+1} - t_n$ itself), tends to a finite, positive limit as $n \to \infty$.

In order to simplify the notations, let the units of length on the x- and t-axes be so chosen that the integration constant (20) (which is positive, since (20 bis) is excluded) and the constant ω occurring in (2) become 1. Since, by (3),

$$\omega^2 + \phi(t) = \omega^2 + o(1) = 1 + o(1) \text{ as } t \to \infty,$$

"the nodes of $x(t)$ are majorized and minorized," in Sturm's sense, by the nodes of the non-trivial solutions of the two linear oscillators

$$y'' + (1 \pm \epsilon)y = 0,$$

where ϵ is an arbitrarily fixed positive constant and t is sufficiently large (viz., $T < t < \infty$, where T depends only on ϵ). Since the n-th half-waves of the solutions $y(t) \not\equiv 0$ of these linear oscillators are of length $\pi/(1 \pm \epsilon)^{\frac{1}{2}}$ for every n, it follows that every solution $x(t) \not\equiv 0$ of (2) has a sequence of zeros, and that, if $t_1, t_2, \cdots$, where $t_n < t_{n+1}$, denotes the sequence of all these zeros (when $x(t)$ is fixed), then

$$t_{n+1} = t_n + \pi + o(1) \text{ as } n \to \infty.$$

Since $x(t)$ vanishes at $t = t_n$ and at $t = t_{n+1}$, a partial integration gives

$$\int_{t_n}^{t_{n+1}} -x''(t)x(t)\,dt = \int_{t_n}^{t_{n+1}} x'^2(t)\,dt.$$

But (2), where $\omega = 1$, shows that

$$-x''(t)x(t) = x^2(t)[1 + \phi(t)] = x^2(t) + o(1) \text{ as } t \to \infty,$$

by (3) and by the first of the estimates (17). Hence, the integral on the left becomes

$$\int_{t_n}^{t_{n+1}} \{x^2(t)\,dt + o(1)\}dt$$

as $n \to \infty$. On the other hand, since the constant (20) has been chosen to be 1, it is seen from (18) and (1), where $\omega = 1$, that, as $t \to \infty$,

$$x'^2(t) + x^2(t) = 2 + o(1),$$

again by (3) and by the first of the relations (17). Hence, the integral on the right is

$$\int_{t_n}^{t_{n+1}} \{2 - x^2(t) + o(1)\}dt.$$

Consequently,

$$\int_{t_n}^{t_{n+1}} x^2(t)\,dt + \int_{t_n}^{t_{n+1}} o(1)\,dt = \int_{t_n}^{t_{n+1}} 2dt - \int_{t_n}^{t_{n+1}} x^2(t)\,dt.$$

In view of $t_{n+1} - t_n \to \pi$, where $n \to \infty$, this means that

$$\int_{t_n}^{t_{n+1}} x^2(t)\,dt = \pi + \int_{t_n}^{t_{n+1}} o(1)\,dt + o(1) = \pi + o(1)$$

and completes, therefore, the proof of (ii bis).

6. In order to abbreviate the formulae, the following proof of the *asymptotic* assertions of (i) will be carried out under the unessential (since *local*) assumption that $\phi(t)$ has a continuous derivative, $\phi'(t)$. Without this simplification, the proof ought to work with Stieltjes differentials or, rather, with partial integrations which correspond to the interpretation, (22 bis), of (22) in the above proof of (ii).

Clearly, the assertion of (i) is equivalent to the statement that every solution of

$$x'' + (1 + \phi(t))x = 0 \tag{24}$$

is a superposition of the real and imaginary parts of a (particular) complex-valued solution of the form

$$x(t) = y(t) + \epsilon(t), \tag{25}$$

where $y(t)$ is an abbreviation for

$$y(t) = \exp i(t + \tfrac{1}{2}\int^t \phi(s)\,ds) \tag{26}$$

and $\epsilon(t)$ is some function which, together with its derivative, satisfies (5). Correspondingly, (i) will be proved along the lines of the above proof of (i*); that is, the function (26), which becomes a solution, (25), only after the addition of a correction term, will first be substituted into the differential operator on the left of (24), leading to an error in (24). The deviation from 0 of the value of the differential operator for $y(t)$ will then be estimated so as to lead to (5). This will be accomplished by an appropriate application of Lagrange's principle of "varied constants."

First, the derivative of the function (26) is

$$y' = i(1 + \tfrac{1}{2}\phi)y.$$

If this representation of y' is differentiated (and, by the assumption made before (24), it *can* be differentiated), it follows that

$$y'' = i^2(1 + \tfrac{1}{2}\phi)^2 y + \tfrac{1}{2}i\phi' y.$$

Clearly, the last relation can be written in the form

$$y'' + (1 + \phi(t) + \psi(t))y = 0, \tag{28}$$

if ψ denotes the (complex-valued) function defined by

$$2\psi(t) = \tfrac{1}{2}\phi^2(t) - i\phi'(t). \tag{29}$$

By the assumption made before (24), the function (29) is continuous. Furthermore, by the assumptions of (i), the function $\phi(t)$ belongs to both classes (L^0), (L^2); that is, both (3_0) and (3_2) are satisfied. But (3_0) means that the second, and (3_2) that the first, term on the right of (29) is of class (L^1). Accordingly, $\psi(t)$ is a (complex-valued) continuous function of class (L^1).

Since (26) is a solution of (28), and since what (i) claims for (24) is (25) with (5), that is,

$$x(t) - y(t) \to 0 \text{ and } x'(t) - y'(t) \to 0 \text{ as } t \to \infty, \tag{30}$$

it is clear that (i) will be proved if it is shown that (30) represents a one-to-one correspondence between the solutions, $x(t)$, of (24) and the solutions, $y(t)$, of (28).

For the sake of short formulations, it will be convenient to use the following manner of speaking:

(iii^0) *A continuous (possibly complex-valued) function, $f(t)$, defined for large positive t, will be called of class (*) if*

$$x(t) = O(1) \text{ and } x'(t) = O(1) \text{ as } t \to \infty$$

holds for every solution of the differential equation $x'' + f(t)x = 0$.

In terms of this definition, all that remains to be proved is contained in the following lemma:

(iii) *If $f(t)$ and $g(t)$ are continuous (possibly complex-valued) functions defined for large positive t in such a way that $f(t)$ is of class (*) and $g(t) - f(t)$ of class (L^1), then (30) represents a one-to-one correspondence between the solutions of*

$$x'' + f(t)x = 0 \tag{31}$$

and the solutions of

$$y'' + g(t)y = 0. \tag{32}$$

If (iii) is granted, the proof of (i) can be completed by identifying (24) with (31), and (28) with (32). In fact, $g(t) - f(t)$ then becomes the

function $\psi(t)$ which, as verified after (29), is of class (L^1). Hence, it is sufficient to ascertain that also the other assumption of (iii), that requiring that $f(t)$ be of class (*), is satisfied. But $f(t)$ now is the function $1+\phi(t)$ which, in view of (iii°) and of the by-product, (17), of the Corollary of (ii), is of class (*), since, by the assumptions of (i), the function $\phi(t)$ is of class (L^0).

7. What remains to be ascertained, viz., the truth of (iii), is a particular case of a general theorem, appearing elsewhere, which concerns arbitrary systems of linear differential equations of first order. Due to the (formally) self-adjoint character of the (possibly complex) differential equations (32), (31), the proof of (iii) itself affords certain simplifications (and leads, in addition, to a refinement of (iii); cf. (iv) below). It will be given in a form which, instead of the classical procedure of successive aproximations, depends only on an adaptation of the trivial estimates used in the proof of (ii).

Let $x=u(t)$ and $x=v(t)$ be two linearly independent solutions of (31). Since their Wronskian, $u(t)v'(t)-v(t)u'(t)$, is a non-vanishing constant (Abel), it can be assumed to be 1. Then

$$X^{-1}(t)=\begin{pmatrix} v'(t) & -v(t) \\ -u'(t) & u(t) \end{pmatrix}, \quad \text{if} \quad X(t)=\begin{pmatrix} u(t) & v(t) \\ u'(t) & v'(t) \end{pmatrix}.$$

Hence, if

$$A(t)=\begin{pmatrix} 0 & 1 \\ -f(t) & 0 \end{pmatrix} \text{ and } \quad B(t)=\begin{pmatrix} 0 & 1 \\ -g(t) & 0 \end{pmatrix},$$

then

$$X^{-1}(B-A)X=(f-g)\begin{pmatrix} -uv & -v^2 \\ u^2 & uv \end{pmatrix}.$$

But (31), (32) can be written as $\xi'=A(t)\xi$, $\eta'=B(t)\eta$, if ξ, η denote the binary vectors the first and second components of which are x, x' and y, y', respectively. On the other hand, Lagrange's rule for the variation of constants (which can, of course, be verified by direct substitution of the matrices and vectors involved) states that, by virtue of the Wronskian transformation, $X(t)$, of the binary system $\xi'=A(t)\xi$, the binary system $\eta'=B(t)\eta$ is equivalent to

$$\zeta'=C(t)\zeta, \qquad C=X^{-1}(B-A)X,$$

if $\zeta = \zeta(t)$ denotes the binary vector defined by $\zeta = X^{-1}(t)\eta$, where $\eta = \eta(t)$ is an *arbitrary* solution of $\eta' = B(t)\eta$.

Let $p = p(t)$ denote the first, and $q = q(t)$ the second, component of $\zeta = \zeta(t)$. Then the last two formula lines mean that

$$(33) \qquad p' = (g - f)(uvp + v^2 q), \qquad q' = (f - g)(u^2 p + uvq).$$

On the other hand, $X(t)$ having been defined to be the Wronskian matrix of $u = u(t)$ and $v = v(t)$, the definition, $\zeta = X^{-1}(t)\eta$, of ζ means that

$$(34) \qquad y = up + vq, \qquad y' = u'p + v'q.$$

Since $x = u(t)$ and $x = v(t)$, being solutions of (31), are assumed to be bounded *along with their derivatives*, $u'(t)$ *and* $v'(t)$, and since the determinant of the linear transformation, (34), of (p, q) into (y, y') is the Wronskian, which is 1 for every t, finally, since every solution, $x = x(t)$, of (31) is a unique superposition of the two particular solutions, $x = u(t)$ and $x = v(t)$, which have been selected for $X(t)$, it is clear that the assertion of (iii) is equivalent to the following statement: If $p = p(t)$, $q = q(t)$ is any solution of the binary linear system (33) (in which u and v are *given* functions of t, as are f and g), then $p(t)$, $q(t)$ tend, as $t \to \infty$, to finite limits,

$$(35) \qquad p(\infty), q(\infty),$$

both of which vanish only for the trivial solution,

$$(35 \text{ bis}) \qquad p(t) \equiv 0 \equiv q(t),$$

of (33).

8. In order to prove the existence of the finite limits (35) for every solution of (33), exclude the trivial solution (35 bis). Then, by the uniqueness theorem of linear systems of differential equations, $p(t)$ and $q(t)$ cannot vanish simultaneously. Hence, if $r(t)$ denotes the function

$$(36) \qquad r(t) = (|p(t)|^2 + |q(t)|^2)^{\frac{1}{2}}$$

(which is positive, whereas $p(t)$ and $q(t)$ can be complex-valued), then the derivative $r'(t)$ exists ($\neq \pm \infty$), since, according to (33), both $p'(t)$ and $q'(t)$ exist.

On adding the relations which result if the first of the equations (33) is multiplied by p and the second by q, and then taking the analogous com-

binations belonging to complex conjugates, one readily sees from (36) that, since $2\,|\,u(t)v(t)\,| \leqq |\,u(t)\,|^2 + |\,v(t)\,|^2$,

$$|\,r(t)r'(t)\,| \leqq \text{Const.}\,|\,f(t) - g(t)\,|\,r^2(t)\,(|\,u(t)\,|^2 + |\,v(t)\,|^2)$$

holds for a certain absolute constant. In view of $r(t) > 0$, this inequality implies that, as $t \to \infty$,

$$(37) \qquad |(\log r(t))'\,| = O(1)\,|\,f(t) - g(t)\,|\,(|\,u(t)\,|^2 + |\,v(t)\,|^2).$$

But (iii°) shows that what is required of $f(t)$ in (iii) is that all four functions

$$(38) \qquad u(t), v(t); u'(t), v'(t)$$

be $O(1)$, since $x = u(t)$ and $x = v(t)$ are two (linearly independent) solutions of (30). On the other hand (even if just the first two of the four functions (37 bis) are $O(1)$ as $t \to \infty$), it follows from (37) that

$$(39) \qquad \log r(t) = O(1) \int^t |\,f(s) - g(s)\,|\,ds = O(1),$$

by the assumption made in (iii) for $f(t) - g(t)$. Since (39) implies that $r(t) = O(1)$, it now follows from (36) that both $p(t)$ and $q(t)$ are $O(1)$. Hence, it is seen from (33) that

$$(40) \qquad p'(t) = O(1)\,|\,f(t) - g(t)\,|\,, \qquad q'(t) = O(1)\,|\,f(t) - g(t)\,|\,,$$

since the functions (38) are $O(1)$ (actually, only the first two of the four functions (38) are needed).

Since the integral occurring in (39) is $O(1)$, it follows from (40) that both integrals

$$(41) \qquad \int^\infty p'(t)\,dt, \qquad \int^\infty q'(t)\,dt$$

are absolutely convergent. In particular, they are convergent. But their convergence proves the existence of the finite limits (35).

Finally, at least one of the two values (35) is distinct from 0. For, if the contrary is assumed, it follows from (36) that $r(t) \to 0$ as $t \to \infty$. Since this implies that $\log r(t) \neq O(1)$, it contradicts (39).

9. This proves (iii), hence (i) as well, and so, since (ii) has been proved earlier, all of the statements made above.

The method applied leads somewhat further than (iii), In order to

express the resulting refinement conveniently, use will be made of the following definition:

(iv^0) *Let a continuous function $f(t)$, defined for large positive t, be called of class* [*] *if $x(t) = O(1)$ holds for every solution of $x'' + f(t)x = 0$.*

Thus, in contrast to (iv^0), it is now not required that $x'(t) = O(1)$. Correspondingly, the generalization alluded to before is as follows:

(iv) *In* (iii), *the assumption that $f(t)$ be of class* (*) *can be relaxed to the assumption that $f(t)$ be of class* [*], *if* (30) *in the assertion of* (iii) *is relaxed to*

$$x(t) - y(t) \to 0 \text{ as } t \to \infty. \tag{30 bis}$$

In accordance with the remark made after (iv^0), only the first of the two relations (30) is involved in the one-to-one correspondence claimed in (iv). Correspondingly, the truth of (iv) can be seen as follows:

In the preceding proof of (iii), the assumption requiring $f(t)$ to be of class (*) has been used three times, the first time after (34), the second time before (39) and the last time after (40). However, the estimate $O(1)$ has been used only for the first two, rather than for all four, of the functions (38). But the assumption, (iv^0), of (iv) means that the first two of the functions (38) are $O(1)$. Hence, the proof of (iii) makes it clear that, under the assumptions of (iv), both components, $p(t)$ and $q(t)$, of every solution of (33) tend to finite limits, (35), both of which vanish only in the case (35 bis). Hence, if $p(\infty) = c_1$ and $q(\infty) = c_2$, the first of the relations (34) gives

$$y(t) = u(t)(c_1 + o(1)) + v(t)(c_2 + o(1)).$$

Since $u(t) = O(1)$ and $v(t) = O(1)$, this can be written in the form

$$y(t) = x(t) + o(1), \text{ where } x(t) = c_1 u(t) + c_2 v(t). \tag{42}$$

In order to complete the proof of (iv), all that remains to be ascertained is that the *single-valued* correspondence established by (42) is a *schlicht* correspondence between the solutions of (31) and (32). In other words, it is sufficient to assure that, by virtue of (42), where $(u(t), v(t))$ denotes a fixed pair of two linearly independent solutions of (31), the same pair of integration constants, (c_1, c_2), cannot belong to two distinct solutions, say $y = y^*(t)$ $y = y^{**}(t)$, of 32). But, if this were not true, $y = y^* - y^{**}$ would be a solution $y(t) \not\equiv 0$ of (32) satisfying (42) with $c_1 = 0 = c_2$. This contains, however, a contradiction. In fact, $c_1 = 0 = c_2$ means that $p(\infty) = 0$

$= q(\infty)$, hence $p(t) \equiv 0 \equiv q(t)$, and so $y(t) \equiv 0$, by the first of the relations (34). This contradiction completes the proof of (iv).

It remains to be seen that (iv) actually is more general than (iii):

(iv*) *The coefficient function, $f(t)$, of* (31) *can be of class* (*) *without being of class* [*]. *But this cannot happen if* $f(t) = O(1)$.

In particular, it cannot happen for (2) if $\phi(t)$ satisfies (3).

First, two differentiations show that

$$x(t) = \exp(-t + i \exp 2t)$$

is a solution of (31) when

$$f(t) = -1 + 4 \exp 4t.$$

Since this $f(t)$ is real, but $x(t)$ is not, the general solution of (31) is the superposition of the real and imaginary parts of $x(t)$. It is also clear that $x(t) = O(1)$ but $x'(t) \neq O(1)$; in fact, neither the real nor the imaginary part of $x'(t)$ is $O(1)$. This proves more than what is claimed by the first part of (iv*).

In order to prove the second part of (iv*), it is sufficient to show that, if $f(t) = O(1)$ and $x(t) = O(1)$, then $x'(t) = O(1)$ holds by virtue of (31). But the two O-assumptions imply, by (30), that $x'' = O(1)$, and $x(t) = O(1)$ *and* $x''(t) = O(1)$ always imply that $x'(t) = O(1)$ (Hadamard).

9 bis. It may be mentioned that a trivial modification of the above proof of (iii) and (iv) leads to an easy proof of Weyl's central results ([3], pp. 223-231) concerning his *Grenzkreisfall* (results which replace (43) below by

$$f(t) - g(t) = \text{const.} \tag{43*}$$

but, as seen from his proofs, apply under the more general assumption (43) also). In particular, the Corollary of the following theorem is nothing but [the obvious extension, from (43*) to (43), of] Weyl's main theorem ([3], p. 238) concerning his alternative of the *Grenzkreisfall* and the *Grenzpunktfall*:

(iv bis) *If $f(t)$ and $g(t)$ are (continuous, possibly complex-valued) functions for which*

$$f(t) - g(t) = O(1) \tag{43}$$

holds as $t \to \infty$, and if (31) *has two linearly independent solutions, say $x = u(t)$ and $x = v(t)$, both of which are of class (L^2), then there belongs*

to every solution, $y(t)$, of (32), a pair of (continuous, bounded) functions, say $p(t)$ and $q(t)$, satisfying

$$y(t) = p(t)u(t) + q(t)v(t)$$

and

$$|p(t)|^2 + |q(t)|^2 \to C \text{ as } t \to \infty,$$

where the integration constant $C = C_{y(t)} \geqq 0$ is 0 only when $y(t) \equiv 0$.

COROLLARY. The assumption (43) implies that all solutions of (31) are of class (L^2) if (and/or only if) all solutions of (32) are of class (L^2); simply because the last two formula lines imply that

$$y(t) = O(1)u(t) + O(1)v(t)$$

as $t \to \infty$.

In order to prove the Tauberian "o-refinement," (iv bis), of the latter "O-theorem," let the trivial solution, (35 bis), of (33) be excluded. Then (37) is applicable. But (37) and (43) imply that

$$\int^{\infty} |(\log r(t))'| \, dt \leqq \text{const.} \int^{\infty} (|u(t)|^2 + |v(t)|^2) dt,$$

and the integral on the right of this inequality is convergent, since $u(t)$ and $v(t)$ are suposed to be of class (L^2). Accordingly, $(\log r(t))'$ is absolutely integrable, hence integrable, over the half-line. In view of (36), the balance of the proof of (iv bis) is the same as that of (iii) or (iv) was.

Incidentally, (iv bis) can be thought of as a limiting case, $(\lambda; \lambda/(\lambda - 1)) = (1; \infty)$, of the following remark:

(iv* bis) *If $f(t) - g(t)$ is of class (L^μ), where $\mu > 1$, and if (31) has two linearly independent solutions, say $x = u(t)$ and $x = v(t)$, the squares of which are of class (L^λ), where $\lambda = \mu/(\mu - 1)$, then there belong to every solution $y(t) \not\equiv 0$ of (32) a positive constant, C, and a pair of functions, $p(t)$ and $q(t)$, in terms of which $y(t)$ is representable in the form claimed by* (iv bis).

In fact, the estimate (37) and the assumptions of (iv*) imply, by Hölder's inequality, the convergence of the integral on the right of the last formula line, and so (iv* bis) follows in the same way as (iv bis) did.

More useful than (iv* bis) is its limiting case $(\lambda; \lambda/(\lambda - 1)) = (\infty; 1)$:

(iv_0) *If $f(t) - g(t)$ is of class $(L) = (L^1)$, and if (31) has two linearly independent solutions, say $x = u(t)$ and $x = v(t)$, which are $O(1)$*

as $t \to \infty$, *then there belong to every solution* $y(t) \not\equiv 0$ *of* (32) *a positive constant,* C, *and a pair of functions,* $p(t)$ *and* $q(t)$, *in terms of which* $y(t)$ *is representable in the form claimed by* (iv bis).

In fact, this dual, $(\lambda;\mu) = (\infty;1)$, of (iv bis), where $(\lambda;\mu) = (1;\infty)$, and even more than this dual, is contained in (iv), since, in view of the definition, (iv^0), preceding (iv), the assumptions of the last theorem, (iv_0), are precisely those of (iv) itself.

10. After this deviation from the direction of (i), it will now be shown that the assertions of (i bis) need not hold if (3_0), the first of the assumptions of (i bis), is omitted; not even if (3), which then becomes a condition not implied by the remaining assumptions of (i bis), is satisfied. More than this is contained in the following negation:

(i* bis) *The convergence of both integrals* (7) *and the assumption* (3) *together do not prevent for* (2) *a solution* $x(t) \not\equiv O(1)$.

Conversely, the three assumptions of this negation, (i* bis), do not prevent for (2) a solution $x(t) \not\equiv 0$ which (instead of being, as in (i* bis), "large") is "small" as $t \to \infty$. In fact, the proof of (i* bis) will be such as to supply, for every $N > 0$, a $\phi(t)$ which satisfies the three assumptions of (i* bis) but is such that (2) admits of a solution $x(t) \not\equiv 0$ which is $O(t^{-N})$ as $t \to \infty$. The existence of functions $\phi(t)$ of either of these types can readily be concluded from the following rule of construction:

If $\chi = \chi(t)$ *is any function having a continuous derivative,* χ', *and if* $\phi = \phi(t)$ *denotes the* (*continuous*) *function*

$$-\phi = \chi^2 \cos^2 t + \chi' \cos t - 3\chi \sin t, \tag{44}$$

then

$$x(t) = \exp\Big(\int^t \chi(s) \cos s \, ds\Big) \cos t \tag{45}$$

is a solution of

$$x'' + (1 + \phi(t))x = 0$$

(that is, of (2), where $\omega = 1$).

For similar purposes, this rule has been deduced in [4], pp. 394-395 [but with an error in sign, the sum of the second and third terms on the right of the above definition, (44), of ϕ, viz., the sum

$$(\chi' \cos t - \chi \sin t) - 2\chi \sin t = \chi' \cos t - 3\chi \sin t,$$

having been given (loc. cit., middle of page 395, where χ is denoted by g), erroneously, as

$$(\chi' \cos t - \chi \sin t) + 2\chi \sin t = \chi' \cos t - \chi \sin t,$$

which, however, does not affect the construction of the possibilities described there].

Remark. If ϕ, rather than χ, is given, then (44), instead of being the definition of ϕ, is *a (singular) Riccati equation* for χ.

Let the above rule be applied to

$$\chi(t) = t^{-1} \cos t \tag{46}$$

(where, in order to exclude the singularity of (46) at $t = 0$, the half-line is, e.g., $1 \leqq t < \infty$). It is seen from (46) that $\chi'(t)$ is $-t^{-1} \sin t + O(t^{-2})$ as $t \to \infty$. Hence, from (44) and (46),

$$-\phi(t) = O(t^{-1})^2 - t^{-1} \sin t \cos t + O(t^{-2}) - 3t^{-1} \cos t \sin t,$$

which means that

$$\phi(t) = 2t^{-1} \sin 2t + O(t^{-2}). \tag{47}$$

On the other hand, since

$$\int^t s^{-1} \cos^2 s \, ds \sim \theta \log t \text{ as } t \to \infty,$$

where θ is a positive constant, it is seen from (46) that the solution (45) is of the form

$$x(t) = \exp(\theta \log t + o(\log t)) \cos t = t^{\theta + o(1)} \cos t, \tag{48}$$

hence $x(t) \neq O(1)$. Since, as shown by (47), both integrals (7) are convergent, and (3) is satisfied by the present $\phi(t)$, the proof of (i* bis) is complete.

It is also seen that, in order to obtain a $\phi(t)$ belonging to the arbitrary exponent, $-N$, mentioned after (i* bis), it is sufficient to multiply (46) by an arbitrary constant. In fact, (44) shows that (47) then becomes multiplied by a constant, and so all three conditions of (i* bis) remain satisfied. On the other hand, the absolute exponent, θ, occurring in (48) becomes replaced by $c\theta$, where c is an arbitrary constant; so that (45) is a solution $x(t) \not\equiv 0$ which is $O(t^{-N})$ as $t \to \infty$, if c is chosen to be less than $-N/\theta$.

Appendix.

In view of (ii), the negation (i*), in which the assumption is the same as in (ii), is disappointing. In fact, one would like to have *some* asymptotic formula, that is, something like (4), rather than just the estimates (17), (19), even though just an (L^0)-condition on the coefficient function is made.

It turns out that such an asymptotic formula, having a structure similar to, but different from, that of (4), actually exists. In fact, the true theorem is as follows:

If $\phi(t)$ is of class (L^0), then there belongs to every (real-valued) solution $x(t) \not\equiv 0$ of (2) a unique pair of positive integration constants, say a and α $(\leqq 2\pi)$, such that

$$x(t) = a\cos\left(\alpha + \int^t \{\omega^2 + \phi(s)\}^{\frac{1}{2}}\, ds\right) + \epsilon(t),$$

where the remainder term, $\epsilon(t) = \epsilon_{a\alpha}(t)$, and its derivative satisfy (5).

This means that the (L^2)-assumption of (i) becomes superfluous if the assertion, (4), of (i) is replaced by the last formula line. The integral occurring in it is

$$\omega \int^t (1 + \phi(s)/\omega^2)^{\frac{1}{2}}\, ds = \omega \int^t (1 + \tfrac{1}{2}\phi(s)/\omega^2 - \cdots)ds,$$

which, approximately, is

$$\omega \int^t (1 + \tfrac{1}{2}\phi(s)/\omega^2)ds = \omega t + \tfrac{1}{2}\int^t \phi(s)\, ds/\omega,$$

the corresponding phase in (4). But the error in the fluctuations, which is thus introduced by the neglect of the higher terms of the binomial expansion, can affect the asymptotic variation of the phase in such a fashion that the approximation becomes illegitimate without the (L^2)-assumption of (i). This makes (i*) understandable indeed.

Clearly, the theorem to be proved can be formulated as follows: If $\phi(t)$ is of class (L^0), then there belongs to every solution, $x(t)$, of (2) a unique integration constant, say $c = a + ib$, satisfying

$$x(t) - cy(t) \to 0 \text{ and } x'(t) - cy'(t) \to 0 \text{ as } t \to 0,$$

where

$$y(t) = \exp i \int^t (\omega^2 + \phi(s))^{\frac{1}{2}} ds.$$

But, if (2) is identified with (31), then (ii_0) and (iii^0) show that the coefficient function, $f(t)$, of (31) is of class (*). It follows therefore from (iii) that the proof will be complete if it is shown that the function defined by the last formula line satisfies a differential equation, (32), whose coefficient function, $g(t)$, is such as to make the deviation $f(t) - g(t)$ a function of class (L^1).

In the deduction of this differential equation, it can, for the reasons explained in connection with (22 bis), be assumed that $\phi(t)$ has a continuous derivative, $\phi'(t)$. Then, since a differentiation of the last formula line gives

$$y' = i(\omega^2 + \phi)^{\frac{1}{2}} y,$$

one more differentiation shows that

$$y'' = \tfrac{1}{2} i(\omega^2 + \phi)^{-\frac{1}{2}} \phi' y - (\omega^2 + \phi) y.$$

This means that $y(t)$ becomes a solution of (32) if $g(t)$ is defined by

$$g(t) = \omega^2 + \phi(t) - \tfrac{1}{2} i(\omega^2 + \phi(t))^{-\frac{1}{2}} \phi'(t).$$

Since (31) is represented by (2), it follows that

$$\int^{\infty} |\, g(t) - f(t) \,|\, dt = \int^{\infty} |\, -\tfrac{1}{2} i(\omega^2 + \phi(t))^{-\frac{1}{2}} \,|\, |\, d\phi(t) \,|\,.$$

But $\phi(t)$ is supposed to be of class (L^0), that is, $\phi(t)$ satisfies (3_0) and (3). Consequently,

$$\int^{\infty} |\, g(t) - f(t) \,|\, dt \leqq \int^{\infty} \text{const.} \,|\, d\phi(t) \,| < \infty.$$

Since this means that $g(t) - f(t)$ is of class (L^1), the proof is complete.

It is clear that, if $\omega(t)$ is a positive, continuous function which tends, as $t \to \infty$, to a positive limit, $\omega(\infty)$, then condition (3_0) is satisfied by the difference $\phi(t) = \omega(t) - \omega(\infty)$ if and only if it is satisfied by $\phi(t) = \omega^2(t) - \omega^2(\infty)$, the difference of the squares. Hence, the preceding theorem can be restated as follows:

If $\lambda(x)$, where $0 \leqq x < \infty$, is a positive, continuous function satisfying

$$\int^{\infty} |\, d\lambda(x) \,| < \infty \text{ and } \lambda(\infty) > 0,$$

then there belongs to every solution, $\psi = \psi(x)$, of the corresponding wave equation,

$$d^2\psi/dx^2 + \{2\pi/\lambda(x)\}^2\psi = 0,$$

a unique integration constant, $c = a + ib$, *such that the difference*

$$\psi(x) - c \exp\,(2\pi i \int_0^x \{\lambda(s)\}^{-1}\, ds),$$

as well as the derivative of this difference, tends to 0 *as* $x \to \infty$.

In fact, the restrictions imposed on $\lambda(x)$ are equivalent to

$$\int^{\infty} |\, d\omega(x)| < \infty \text{ and } \omega(\infty) > 0,$$

where $\omega(x) = 2\pi/\lambda(x)$.

Applications of this theorem to the Hellinger decomposition of the spectral form of the differential equation (in case of a fixed boundary condition,

$$[\alpha\psi(x) + \beta d\psi(x)/dx]_{x=0} = 0, \qquad \alpha^2 + \beta^2 > 0,$$

assigned for a given $x = 0$) and, in particular, to problems in wave mechanics (Kramers) will be given elsewhere.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] A. Kneser, "Untersuchung und asymptotische Darstellung der Integrale gewisser linearer Differentialgleichungen bei grossen reellen Werthen des Arguments (Zweiter Aufsatz)," *Journal für die reine und angewandte Mathematik*, vol. 117 (1897), pp. 72-103.

[2] O. Perron, "Ueber ein vermeintliches Stabilitätskriterium," *Göttinger Nachrichten*, 1930, pp. 28-29.

[3] H. Weyl, "Ueber gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen," *Mathematische Annalen*, vol. 68 (1910), pp. 222-269.

[4] A. Wintner, "The adiabatic linear oscillator," *American Journal of Mathematics*, vol. 68 (1946), pp. 385-397.

POLYNOMIAL LEAST SQUARE APPROXIMATIONS.*

By D. C. Lewis.

1. Introduction. Suppose that we have a (not necessarily bounded) linear functional, whose domain is a class of sufficiently regular functions, $f(x)$, and whose range is the class of polynomials, $P_n(x)$, of degree $\leqq n$. Suppose furthermore that this functional is such that $P_n(x) \equiv f(x)$ whenever $f(x)$ is itself a polynomial of degree $\leqq n$. Functionals of this type are commonly used for obtaining polynomial "approximations," $P_n(x)$, to functions, $f(x)$. Elementary examples are afforded by the Lagrange interpolation polynomials and by the polynomials obtained by taking the first $n+1$ terms of Taylor's series or the first $n+1$ terms in an expansion in a series of orthogonal polynomials with respect to an arbitrary weight function. There are many other examples. We accordingly refer to the difference, $f(x) - P_n(x)$, as the "remainder."

The object of this paper is to obtain a simple explicit expression for this remainder in terms of the $(m+1)$-th derivative of $f(x)$ $(m \leqq n)$ and certain other elements independent of $f(x)$ but dependent upon the particular functional under consideration.

Our results are not applicable to the most general functional of the type described above but only to those which involve, in a certain general way, the principle of least squares, or, more generally, those expressible in the manner indicated by equation (3) below. Even so, our results are sufficiently general to present a unified approach to approximations of the Lagrange or Taylor type [1] and to those of the Legendre or Tchebichef type [2] as well as to many other previously uninvestigated types. After reading the paper, the reader will recognize that the Lagrange or Taylor type of approximation is obtained in the cases when $\alpha_0(x), \alpha_1(x), \cdots, \alpha_p(x)$, introduced in the next section, are step functions with just the right discontinuities, totalling $n+1$ in number, to determine $P_n(x)$ uniquely so as to minimize the quantity S of

* Received July 28, 1946.

[1] Cf. G. D. Birkhoff, "General mean value and remainder theorems with applications to mechanical differentiation and quadrature," *Transactions of the American Mathematical Society*, vol. 7 (1906), pp. 107-136.

[2] Cf. G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. 23 (1939).

equation (1), this minimum value being zero. On the other hand the Legendre or Tchebichef type of approximation is characterized by the fact that $\alpha_0(x)$ is a suitable function with an infinite number of points of increase while $p = 0$. The minimum value of S is usually *not* zero. The expression for the remainder is valid outside as well as inside the interval of orthogonality.

2. Specification of the least squares problem. Let $\alpha_0(x)$, $\alpha_1(x)$, $\alpha_2(x)$, $\cdots, \alpha_p(x)$ be $p+1$ monotonic non-decreasing functions defined for $a \leqq x \leqq b$. Let $f(x)$ be a function of class C^{p-1} over an interval $A \leqq x \leqq B$, where $A \leqq a < b \leqq B$. We suppose also that the p-th derivative $f^{(p)}(x)$ exists almost everywhere with respect to α_p and is such that the Lebesgue-Stieltjes integral,

$$\int_a^b [f^{(p)}(x)]^2 d\alpha_p(x),$$

exists. A polynomial $P_n(x)$ of degree $\leqq n$ whose coefficients are such as to render

$$S = \sum_{k=0}^{p} \int_a^b [f^{(k)}(x) - P_n{}^{(k)}(x)]^2 d\alpha_k(x) \tag{1}$$

a minimum, is said to be the polynomial of degree n which best fits the function $f(x)$. This "best fitting" is, of course, in the sense of least squares *with respect to* a given set of distributions $\alpha_0(x), \cdots, \alpha_p(x)$, over the interval $a \leqq x \leqq b$. Since this is the only sense in which the expression is used in this paper, we omit in the sequel this detailed specification. Moreover it will be convenient to denote the relationship between a function $f(x)$ and its best fitting polynomial $P_n(x)$ in the following manner:

$$f(x) \sim P_n(x). \tag{2}$$

In general, $P_n(x)$ in (2) is uniquely determined by $f(x)$. This is well known to be the case if $\alpha_0(x)$ has at least $n+1$ points of increase; but it is also true in many other circumstances: for example, when $\alpha_0(x), \alpha_1(x), \cdots, \alpha_q(x)$ each have at least one point of increase, while α_{q+1} has at least $n - q$ points of increase. $(q+1 \leqq p.)$ The general condition is that the quadratic form,

$$\sum_{k=0}^{p} \int_a^b [(d^k/dx^k)(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n)]^2 d\alpha_k(x),$$

in the $n+1$ quantities $a_0, a_1, \cdots, a_n$ should be positive definite. Under this condition the reader will find without difficulty that the $n+1$ linear

equations for the determination of the $n+1$ coefficients of $P_n(x)$ so as to minimize the expression S have a non-vanishing determinant. A further investigation of the solutions of these $n+1$ linear equations reveals the important fact that $P_n(x)$ can be written in the form,

$$P_n(x) = \sum_{k=0}^{p} \int_a^b K_{nk}(x, y) f^{(k)}(y) d\alpha_k(y), \tag{3}$$

where $K_{nk}(x, y)$ are polynomials in both x and y, independent of f. In the sequel we assume that $\alpha_0, \alpha_1, \cdots, \alpha_p$ are such that $P_n(x)$ is thus always uniquely determined by $f(x)$. A simple but important consequence of this is

LEMMA 1. *If $f(x)$ is a polynomial of degree $\leqq n$ and if $f(x) \sim P_n(x)$, then $f(x) \equiv P_n(x)$. In other words, by* (3)

$$Q_\mu(x) = \sum_{k=0}^{p} \int_a^b K_{nk}(x, y) Q_\mu^{(k)}(y) d\alpha_k(y), \tag{4}$$

for any polynomial $Q_\mu(x)$ of degree $\mu \leqq n$.

The proof consists in the remark that the indicated determination of $P_n(x)$ gives S in (1) the value zero, which is clearly a minimum value for the non-negative S.

3. Statement and proof of the main theorem. From now on we assume that $f(x)$ is of class $C^{m-1}(m \geqq p)$ on the interval $A \leqq x \leqq B$ and that the m-th derivative exists almost everywhere, is of bounded variation, and is continuous at $x = b$. It must also be assumed that $f^{(m-1)}(x)$ is *absolutely* continuous. In case $m = p$, we assume that $f^{(m)}(x)$ and $\alpha_p(x)$ have no points of discontinuity in common.

Let $f(x) \sim P_n(x)$. We now state our principal

THEOREM. *There exists a function $G_n^m(x, t)$, independent of $f(x)$, with $p \leqq m \leqq n$, such that*

$$f(x) = P_n(x) + (1/m!) \int_a^b G_n^m(x, t) df^{(m)}(t) + (1/m!) \int_b^x (x - t)^m df^{(m)}(t), \tag{5}$$

for any point x on the interval $A \leqq x \leqq B$. Moreover the function $G_n^m(x, t)$ is explicitly given for each value of t as that polynomial in x of degree $\leqq n$ which best fits the function $g_m(x, t)$ defined as follows:

$$g_m(x, t) = (x - t)^m \quad \text{if } x \leqq t$$
$$g_m(x, t) = 0 \quad \text{if } x > t.$$

We begin our proof by referring to Taylor's theorem with the integral form of the remainder:

$$(6) \qquad f(x) = \sum_{h=0}^{m} (f^{(h)}(b)/h!)(x-b)^h + (1/m!)\int_b^x (x-t)^m df^{(m)}(t).$$

This shows that the identity (5) to be established is equivalent to the following:

$$(7) \qquad P_n(x) = \sum_{h=0}^{m} (f^{(h)}(b)/h!)(x-b)^h - (1/m!)\int_b^b G_n{}^m(x,t)\,df^{(m)}(t).$$

To establish (7), we insert the right hand member of (6), after changing x to y, into (3). If we use Lemma 1, we discover at once that

$$P_n(x) = \sum_{h=0}^{m}(f^{(h)}(b)/h!)(x-b)^h + \sum_{k=0}^{p}\int_a^b K_{nk}(x,y)\,(1/(m-k)!)\int_b^y (y-t)^{m-k}df^{(m)}(t)\,d\alpha_k(y).$$

If we reverse the order of the iterated integrals, we obtain

$$P_n(x) = \sum_{h=0}^{m}(f^{(h)}(b)/h!)(x-b)^h + \sum_{k=0}^{p}(1/(m-k)!)\int_b^a \int_a^t K_{nk}(x,y)\,(y-t)^{m-k}d\alpha_k(y)\,df^{(m)}(t).$$

But this reduces at once to (7), if we make use of the formulas,

$$G_n{}^m(x,t) = \sum_{k=0}^{p}\int_a^b K_{nk}(x,y)\,g_m{}^{(k)}(y,t)\,d\alpha_k(y) = \sum_{k=0}^{p}(m!/(m-k)!)\int_a^t K_{nk}(x,y)\,(y-t)^{m-k}d\alpha_k(y),$$

which are clear consequences of the definition of $G_n{}^m$ and of (3).

This completes the formal proof. The only delicate step of rigor is the reversal of the order of the iterated Stieltjes integrals, which, of course, involves Fubini's theorem.[3] The application of Fubini's theorem to the double integral,

$$\int\int K_{nk}(x,y)\,g_m{}^{(k)}(y,t)\,d\alpha_k(y)\,df^{(m)}(t),$$

over the region, $a \leqq y \leqq b$, $a \leqq t \leqq b$, is, however, immediate when the integrand is continuous. The only case of discontinuity occurs when $k = m = p$, in which case $g_m{}^{(k)}(y,t)$ is not defined at the points of discontinuity, namely the points for which $y = t$. In virtue of the hypothesis

[3] Cf. Stanislaw Saks, *Theory of the Integral* (2nd edition), p. 77.

that $\alpha_p(x)$ and $f^{(m)}(x)$ have no common points of discontinuity, it is easily seen that the set $y = t$ has measure zero with respect to the two dimensional measure function, $m(E) = \int_E \int d\alpha_p(y) df^{(m)}(t)$. Hence, Fubini's theorem can be applied even in this case.

The reason for the hypothesis that $f^{(m)}(x)$ be continuous at $x = b$ and also for the hypothesis that $f^{(m-1)}(x)$ be absolutely continuous comes in the proof of Taylor's Theorem in the form indicated in (6). This formula is established by repeated integration by parts in a well known manner, but the proof requires the assumption that $f^{(m)}(b-0) = f^{(m)}(b)$, if $b < x$, and that $f^{(m)}(b+0) = f^{(m)}(b)$, if $b > x$.[4] It is equally necessary to assume that $f^{(m-1)}(x)$ is an indefinite integral of $f^{(m)}(x)$, which means that $f^{(m-1)}(x)$ must be absolutely continuous. That these conditions are not only required by the proof but are also logically necessary for the validity of (6) may be shown easily with the help of simple counter examples.

4. Interpretation of the main theorem in case $a \leqq x \leqq b$. If $a \leqq x \leqq b$, it is obvious from the definition of $g_m(x,t)$ that (5) can be written in the form,

$$(8) \qquad f(x) - P_n(x) = (1/m!) \int_a^b [G_n{}^m(x,t) - g_m(x,t)] df^{(m)}(t).$$

Now, in its dependence upon x, we have $g_m(x,t) \sim G_n{}^m(x,t)$. Equation (8) thus indicates the manner in which the remainder for the *general* function $f(x)$ is synthesized from the remainder for the *special* function $g_m(x,t)$ and from the values of $f^{(m)}(x)$.

The result embodied in (8) is probably well known in various special cases, although specific references are unknown to the author.

In the fairly general situation where at least one of the α's has infinitely many points of increase, one can, with fixed α's and fixed m, let $n \to \infty$. It is then clear from (8) that, if $G_n{}^m$ is uniformly bounded and approaches $g_m(x,t)$ almost everywhere with respect to $\int |df^{(m)}(t)|$, $P_n(x)$ must approach $f(x)$. Here, of course, m may be taken to be any number $\geqq p$. In case $p = 0$, m may also be taken to be zero, and then $f^{(0)}(x)$, is, of course, $f(x)$ itself. These remarks have obvious connections with well known results on the convergence of series of orthogonal polynomials.

---

[4] Cf. Saks, *loc. cit.*, p. 102.

5. Taylor's remainder theorem as a special case of the main theorem.

Although our main theorem is a consequence of Taylor's theorem with the integral remainder, it is of interest to show that the converse is also true. This will give an indication of the scope of our result by discussing a case not covered in the last paragraph of the preceding section.

Let $p = m = n$. Let c be any point of the interval (a, b) where $f^{(m)}(x)$ is continuous, and let $\alpha_k(x) = 0$ for $a \leqq x \leqq c$, and $\alpha_k(x) = 1$ for $c < x \leqq b$, $k = 0, 1, 2, \cdots, n$. Then the requirement that S be a minimum reduces to the following:

$$P_n^{(k)}(c) = f^{(k)}(c), \text{ for } k = 0, 1, 2, \cdots, n.$$

Hence $P_n(x)$ is the sum of the first $n+1$ terms of the Taylor expansion of of $f(x)$ about the point $x = c$. Similarly, since $g_n(x, t) \sim G_n^n(x, t)$, we have

$$\frac{d^k G_n^n(c, t)}{dx^k} = \frac{d^k g_n(c, t)}{dx^k}, \quad k = 0, 1, 2, \cdots, n.$$

This shows that $G_n^n(x, t) = (x - t)^n$, if $c < t$,

$$G_n^n(x, t) = 0, \text{ if } t < c.$$

Equation (5) can therefore in this particular instance be written as follows:

$$\begin{aligned} f(x) &= P_n(x) + (1/m!) \int_c^b (x - t)^n df^{(m)}(t) \\ &\quad + (1/m!) \int_b^x (x - t)^m df^{(m)}(t) \\ &= \sum_{k=0}^{n} (f^{(k)}(c)/k!)(x - c)^k + (1/m!) \int_c^x (x - t)^n df^{(m)}(t), \end{aligned}$$

which is merely equation (6) with b replaced by c.

UNIVERSITY OF MARYLAND.

A NOTE ON LAPLACE TRANSFORMS OF FUNCTIONS WHOSE SPECTRA ARE CONFINED TO A GIVEN SET.*

By E. K. HAVILAND.

It is known that a necessary and sufficient condition that a function $f(x)$ should be representable in the form

$$\int_0^\infty e^{-xt} d\beta(t),$$

where $\beta(t)$ is bounded and non-decreasing and the integral converges for $0 \leq x < +\infty$, is that $f(x)$ should be completely monotonic in $0 \leq x < +\infty$. If the requirements on $\beta(t)$ are merely that it be non-decreasing and that the integral converge for $0 < x < +\infty$, the corresponding necessary and sufficient condition is that $f(x)$ should be completely monotonic in $0 < x < +\infty$. For these theorems, several proofs are known, of which the earliest appears to be that which is an immediate consequence of Hausdorff's solution of the continuous momentum problem.[1]

The purpose of the present note is to extend the foregoing results on the Laplace transform to the case where the spectrum of the function $\beta(t)$ is confined to a preassigned subset Γ of the interval $[0, +\infty)$, the extension being based on results previously obtained by the present author in connection with the momentum problem.[2]

By the spectrum of the function $\beta(t)$ is meant the set of points τ such that $\beta(t_2) - \beta(t_1) > 0$ for every interval (t_1, t_2) such that $t_1 < \tau < t_2$.

Furthermore, we recall that a function $f(x)$ is defined to be completely monotonic in an interval $(0, +\infty)$, if its p-th derivative satisfies the condition

$$(-1)^p f^{(p)}(x) \geq 0, \qquad (p = 0, 1, 2, \cdots),$$

in the interior of the interval and if the function $f(x)$ is continuous at the

* Received November 27, 1946.

[1] F. Hausdorff, "Summationsmethoden und Momentfolgen. II," *Mathematische Zeitschrift*, vol. 9 (1921), pp. 280-299, especially, pp. 282-287.

[2] E. K. Haviland, "On the momentum problem for distribution functions in more than one dimension," *American Journal of Mathematics*, vol. 58 (1936), pp. 164-168.

end point 0, if the latter is included in the interval. This is equivalent to the infinitely many difference equations [3]

$$(-1)^p \begin{pmatrix} f(t_0) & f(t_1) & \cdots & f(t_p) \\ t_0 & t_1 & \cdots & t_p \end{pmatrix} \geq 0,$$

where the second factor denotes the p-th divided difference of $f(t)$ formed for any points $t_0, t_1, \cdots, t_p$ (≥ 0).

Our principal result is then given by

THEOREM I. *A necessary and sufficient condition that a function $f(x)$ be representable in the form*

$$f(x) = \int_0^\infty e^{-xu} d\beta(u), \tag{1}$$

where $\beta(u)$ is bounded and non-decreasing and has its spectrum confined to the sub-set Γ of $[0, +\infty)$ and the integral converges for $0 \leq x < +\infty$, is that $f(x)$ be completely monotonic in $[0, +\infty)$ and that, if $a_0 + a_1 e^{-u} + \cdots + a_n e^{-nu}$ be any exponential polynomial non-negative on Γ, then the functional value

$$a_0 f(x) + a_1 f(x+1) + \cdots + a_n f(x+n) \geq 0 \tag{2}$$

for all $x \geq 0$.

Proof. If in (1) we make the transformation $e^{-u} = t$ or $u = -\log t$, (u real for $t > 0$), and if we denote the bounded non-decreasing function $-\beta(-\log t)$ by $\alpha(t)$, we may write $f(x)$ in the form

$$f(x) = \int_0^1 t^x d\alpha(t), \tag{3}$$

in which case the spectrum of $\alpha(t)$ is confined to a set C contained in the interval $[0, 1]$, C being the image on the t-axis of the set Γ on the u-axis. Then, if $\phi(t) \geq 0$ is integrable over C with respect to α,

$$\Phi(x) = \int_0^1 t^x \phi(t) d\alpha(t)$$

exists for all $x \geq 0$ and is a continuous function of x there, except perhaps at $x = 0$, at which point a possible right-hand discontinuity can always be

[3] Cf. F. Hausdorff, *ibid.*, p. 284.

avoided by defining[4] $\alpha(0) = \alpha(+0)$, a procedure which, of course, imposes a restriction on $\Phi(0)$.

Moreover, $\Phi(x) \geq 0$ for all x, $(0 \leq x < +\infty)$, and

$$(-1)^k \Phi^{(k)}(x) = (-1)^k \int_0^1 t^x (\log t)^k \phi(t)\, d\alpha(t)$$

exists and is ≥ 0 for all $x > 0$, $(k = 1, 2, \cdots)$, so that $\Phi(x)$ is completely monotonic in $x \geq 0$. In particular, this is true of any polynomial

$$\phi(t) = a_0 + a_1 t + \cdots + a_n t^n, \tag{4}$$

non-negative on C, in which case,

$$\begin{aligned} \Phi(x) &= \int_0^1 t^x \{a_0 + a_1 t + \cdots + a_n t^n\} d\alpha(t) \\ &= a_0 f(x) + a_1 f(x+1) + \cdots + a_n f(x+n) \geq 0. \end{aligned} \tag{5}$$

This completes the proof of the necessity of the conditions stated in Theorem I, and it may be noted that the functional (2) is not only non-negative but completely monotonic.

Suppose, conversely, that the function $f(x)$ is completely monotonic for $x \geq 0$ and that, for every polynomial (4) and every $x \geq 0$ (or, at least, for $x = 0$, which is all that is needed for the present proof), (2) holds. Let $x = 0$ in (2) and put $f(n) = \mu_n$, $(n = 0, 1, 2, \cdots)$. Then to every polynomial (4) non-negative on C there corresponds the non-negative functional value

$$a_0 \mu_0 + a_1 \mu_1 + \cdots + a_n \mu_n.$$

This condition is known[5] to be sufficient for the existence of a bounded non-decreasing function, $\alpha(t)$, whose spectrum is contained in C and which is such that $\mu_n = \int_0^1 t^n d\alpha(t)$.

Since the corresponding moment function $\mu(x) = \int_0^1 t^x d\alpha(t)$ is completely monotonic for $x \geq 0$, the $\mu_n = \mu(n)$ form a completely monotonic sequence and, as $\Sigma (t_n)^{-1} = \Sigma n^{-1}$ diverges, this completely monotonic sequence

---

[4] Cf. F. Hausdorff, *ibid.*, p. 286.

[5] E. K. Haviland, *ibid.*, p. 164.

determines uniquely[6] a completely monotonic function $f(x)$ for which $f(n) = \mu_n$. Consequently, our original function $f(x)$, which was completely monotonic in $x \geqq 0$ by hypothesis, must be identical with $\mu(x)$, i. e., on setting $t = e^{-u}$ and $\beta(u) = -\alpha(e^{-u})$,

$$f(x) = \int_0^\infty e^{-xu} d\beta(u),$$

where $\beta(u)$ is a bounded monotonic function whose spectrum is contained in the preassigned set Γ.

In a similar way, we have

THEOREM II. *A necessary and sufficient condition that a function $f(x)$ be representable in the form (1), where $\beta(u)$ is non-decreasing and has its spectrum confined to the sub-set Γ of $[0, +\infty)$ and the integral converges for $0 < x < +\infty$, is that $f(x)$ be completely monotonic in $0 < x < +\infty$ and that, if $a_0 + a_1 e^{-u} + \cdots + a_n e^{-nu}$ be any exponential polynomial non-negative on Γ, then the functional value (2) be non-negative for all $x > 0$.*

In fact, the necessity of the condition follows exactly as before. As to the sufficiency, we observe that, for any fixed $\xi > 0$, it follows as in the proof of Theorem I that

$$f(\xi + n) = \mu'_n = \int_0^1 t^n d\alpha_\xi(t), \qquad (n = 0, 1, \cdots),$$

uniquely determines the function

$$f(\xi + x) = \mu_\xi(x) = \int_0^1 t^x d\alpha_\xi(t),$$

which is completely monotonic in $x \geq 0$, $\alpha_\xi(t)$ being a bounded non-decreasing function whose spectrum is confined to the set C. On replacing x by $x - \xi$, we obtain

$$f(n) = \int_0^1 t^n \cdot t^{-\xi} d\alpha_\xi(t) = \int_0^1 t^n d\alpha(t), \qquad (n = 0, 1, 2, \cdots)$$

where

$$\alpha(t) = \int_0^t \tau^{-\xi} d\alpha_\xi(\tau),$$

so that, by virtue of the uniqueness of the solution of this momentum problem, $t^{-\xi} d\alpha_\xi(t) = d\alpha(t)$, for all $\xi > 0$. Finally,

$$f(x) = \int_0^1 t^x d\alpha(t)$$

[6] Cf. F. Hausdorff, *ibid.*, p. 284.

for $x > 0$. As the spectrum of $\alpha_\xi(t)$ is confined to the set C for all $\xi > 0$, the same will be true of $\alpha(t)$, save possibly for the addition of the point $t = 0$. But, if we define $\alpha(0) = \alpha(+0)$, where $\alpha(+0)$ may be $-\infty$, this situation will cause no difficulty, inasmuch as a spectrum is necessarily closed. This completes the proof of Theorem II.

We next consider sub-sets Γ of $[0, +\infty)$ for which there can be found a basis of polynomials non-negative on Γ. In particular, let Γ be $[0, 1]$, to which corresponds the interval $C : [e^{-1}, 1]$. Since any polynomial non-negative on the latter interval can be represented (at least in the limit) in the form

$$\Sigma A(1-x)^m(x-e^{-1})^n,$$

where $m, n = 0, 1, 2, \cdots$ and the A's are non-negative,[7] it follows that (4) may here be replaced by

$$(1-t)^m(t-e^{-1})^n = e^{-n}(-1)^n(1-t)^m(1-et)^n$$

$$= e^{-n}(-1)^n \sum_{j=0}^{m} C_j^m(-1)^j \sum_{\nu=0}^{n} C_\nu^n(-1)^\nu e^\nu t^{j+\nu},$$

the C_j^m, C_ν^n being binomial coefficients.

If we neglect the positive factor e^{-n}, the analogue of equation (2) is now

$$(6) \qquad \sum_{j=0}^{m} C_j^m(-1)^j\{(-1)^n \sum_{\nu=0}^{n} C_\nu^n(-1)^\nu e^\nu f(x+j+\nu)\} \geq 0,$$

or

$$(6\text{ bis}) \qquad (-1)^m\Delta_j^m\{\Delta_\nu^n\{e^\nu f(x+j+\nu)\}\} \geq 0,$$

where $m, n = 0, 1, 2, \cdots$ and $x \geq 0$ in the case of Theorem I or $x > 0$ in the case of Theorem II. The latter, for instance, may be stated explicitly as

THEOREM III. *Necessary and sufficient in order that a function $f(x)$ be representable in the form (1), where $\beta(u)$ is a non-decreasing function with spectrum confined to $[0, 1]$ is that*

(i) *the function $f(x)$ be completely monotonic in $x > 0$,*

(ii) *the condition (6) or (6 bis) hold for all $x > 0$.*

---

[7] Cf., e. g., Pólya und Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. II, p. 83, Ex. 49.

The infinitely many difference conditions contained in (6) suggest that possibly (ii) implies (i), the more so because, as is well known,[8]

$$\Delta_k{}^n f(a) = (n!)^{-1} f^{(n)}(\xi), \qquad a < \xi < a + n. \tag{7}$$

That (ii), however, does not imply (i) can be seen from the example

$$f(x) = 2 + \sin(2\pi x),$$

which is positive for all x and has the period 1. On interchanging the order of the summations in (6), we obtain

$$(-1)^{m+n} \sum_{\nu=0}^{n} C_\nu{}^n (-1)^\nu e^\nu \{\Delta_j{}^m f(x + \nu + j)\}. \tag{8}$$

Due to the periodicity of $f(x)$, (8) vanishes for all x if $m \geq 1$, while, if $m = 0$, it reduces to

$$f(x)(-1)^n \sum_{\nu=0}^{n} C_\nu{}^n (-e)^\nu = f(x)\,(e-1)^n > 0,$$

so that in any case condition (ii) is fulfilled. In spite of this, $f(x)$ is not even monotone and hence (i) is not satisfied.

P. Hartman has pointed out to me that condition (ii) of Theorem II, which, in virtue of (7), may be written as

$$(-1)^m \frac{d^n}{dx^n} \{e^x f^{(m)}(x)\} \geq 0,$$

can be replaced by the simpler conditions

(ii*) $f(x) + f'(x)$ is completely monotone for $x > 0$,

or

(ii**) the sequence of numbers $f(x), f'(x), f''(x), \cdots$ is bounded for some fixed $x = x_0 > 0$.

The necessity of these conditions is seen directly, if it be observed that the upper limit of integration in (1) is now 1. As to the sufficiency, (ii*) implies that the sequence $\{|f^{(n)}(x)|\}$, where $f^{(n)} = d^n f/dx^n$, is monotone non-increasing and hence (ii**) is valid. But, if $\beta(t)$ had a point $x > 1$ in its spectrum,

$$(-1)^n f^{(n)}(x) = \int_0^\infty e^{-xt} t^n d\beta(t) \geq (\gamma - \delta)^n \int_{\gamma-\delta}^\infty e^{-xt} d\beta(t),$$

[8] Cf. O. Hölder, "Zur Theorie der trigonometrischen Reihen," *Mathematische Annalen*, vol. 24 (1884), p. 183.

and as δ may be so chosen that $\gamma - \delta > 1$, the last expression would become infinite with n, thus proving the sufficiency of (ii**).

Moreover, Theorem III remains true, if the interval $[0, 1]$ is replaced by any interval $[a, b]$, where $0 \leq a < b < \infty$, and if the set of conditions (i)-(ii) is replaced by either one of the following sets:

(i*) $f(x)e^{ax}$ is completely monotonic for $x > 0$,

(ii*) $bf(x) + f'(x)$ is completely monotonic for $x > 0$

or

(i**) $f(x)$ is completely monotonic for $x > 0$,

(ii**) $f(x) = O(e^{-ax})$, $x \to \infty$,

(iii**) $f^{(n)}(x) = O(b^n)$, $n \to \infty$, for some fixed $x = x_0 > 0$.

THE JOHNS HOPKINS UNIVERSITY.

ON THE REPRESENTATIONS OF THE SYMMETRIC GROUP.*

(Second Paper)

By G. DE B. ROBINSON.

Introduction. The close relationship between the theory of the representations of the full linear group L and that of the symmetric group was first remarked by Schur and has been studied by many authors.[1] In particular, the reduction of the direct product of two irreducible representations of L is given in a theorem of Littlewood and Richardson proved by the author in 1938.[2] Since the substance of the present paper is so closely related to that of SG_1 it has been given the same title and designated 'Second Paper.'

The paper is divided into two parts. In Part I the familiar Young diagram[3] is generalized in the following manner. Consider an irreducible representation $[\alpha]$ of the symmetric group S_l and a representation $[\beta]$ of S_m where $l = m + n$ and $\beta_i \leqq \alpha_i$ for all i. The *right* diagram $[\beta]$ is said to be *contained* in $[\alpha]$, and the symbolic difference $[\alpha] - [\beta]$ is taken to represent the *skew diagram* composed of these nodes of $[\alpha]$ not belonging to $[\beta]$. Just as one speaks of a *standard* right diagram so one can speak of a *standard* skew diagram, and these latter lead to a representation of S_n which is reducible. This representation is called a *skew representation* as compared with the familiar right representation which is irreducible. The irreducible components of $[\alpha] - [\beta]$ are determined in **3**, and some formulae connecting the degrees are developed in **4**.

In Part II this theory is applied to give a new proof of the Murnaghan-Nakayama[4] recursion formula.

The chief purpose of developing the theory in Part I has been to relate the two remarkable papers by Nakayama[4] to the main body of Young's work.

---

* Received June 22, 1946.

[1] For references consult [5]. The recent paper [1] by Stig Comét is of interest, and complete accounts of the background will be found in Littlewood [2] and Weyl [7].

[2] [5]. This paper will be referred to as SG_1.

[3] Young consistently used the word *tableau*. Several recent writers have since favoured the term *diagram* to which usage we conform. It is necessary to have a term for the intersection of a row and column in a diagram. Young spoke of 'place' or 'position,' but *node* seems to be convenient and suggestive.

[4] [4], Part I, p. 107. This paper will be referred to as N_1, Part II as N_2.

The tools developed, particularly formulae 4.3, 4.4 and 4.5, have proved useful and it is hoped to publish the results of their application to the modular theory shortly.

PART I.

A generalization of the Young diagram.

1. Following the notation of SG_1, let us denote by $P_n(L)$ the n-th *power* representation of the full linear group L and write

1.1 $$\Delta_{(\alpha)}(L) = P_{\alpha_1}(L) \times P_{\alpha_2}(L) \times \cdots \times P_{\alpha_h}(L).$$

Since $P_n(L) = \{n\}$, we have also

1.2 $$\Delta_{(\alpha)}(L) = \{\alpha_1\} \times \{\alpha_2\} \times \cdots \times \{\alpha_h\},$$

and in general we may consider the direct product of two such representations

1.3 $$\Delta_{(\alpha)}(L) = \Delta_{(\beta)}(L) \times \Delta_{(\gamma)}(L),$$

where $\Sigma\beta_i = m$, $\Sigma\gamma_i = n$ and $\Sigma\alpha_i = m + n = l$.

In SG_1 the chief interest was in proving Littlewood and Richardson's theorem for the reduction of $\{\beta\} \times \{\gamma\}$. Here we consider the more general problem of the reduction of the direct product 1.3 when $[\alpha] = [1^l]$:

1.4 $$\begin{aligned}\Delta_{(1^l)}(L) &= \Delta_{(1^m)}(L) \times \Delta_{(1^n)}(L) \\ &= \Sigma\{\beta\} \times \{\gamma\} = \Sigma\{\alpha\}.\end{aligned}$$

How many times will a given irreducible component $\{\alpha\}$ appear in 1.4? Certainly, its coefficient will receive contributions from many $\{\beta\} \times \{\gamma\}$. To find such contributions one has recourse to the Littlewood and Richardson rule which we restate here as follows:

1.5 LR_1: Take the diagram $[\beta]$ intact and add to it the symbols of the first row of $[\gamma]$. These may be added to one row of $[\beta]$, or the symbols may be divided without disturbing their order into any number of sets, the first set being added to one row of $[\beta]$, the second to a subsequent row, and so on. After the addition no row must contain more symbols than a preceding row, and no two added symbols may appear in the same column.

Next add the second row of $[\gamma]$ according to the same rules,

followed by the remaining rows in succession until all the symbols of $[\gamma]$ have been used.

LR_2: These additions shall be such that each symbol of a given row of $[\gamma]$ in the compound diagram must appear in a later row than the symbol on the same column from a preceding row of $[\gamma]$.

A necessary condition for $\{\alpha\}$ to be an irreducible component of $\{\beta\} \times \{\gamma\}$ is that the diagram $[\beta]$ be *contained in* the diagram $[\alpha]$ in the sense that $\beta_i \leq \alpha_i$ for all i. Since

1.6 $$\{\beta\} \times \{\gamma\} = \{\gamma\} \times \{\beta\},$$

the same holds for $[\gamma]$.

Since $\{\beta\}$ appears in $\Delta_{(1^m)}(L)$ with frequency f_β, the frequency of appearance of $\{\alpha\}$ on the right hand side of 1.4 is given by

$$\sum_\beta f_\beta \sum_\gamma {}_\alpha\lambda_\beta{}^\gamma f_\gamma;$$

${}_\alpha\lambda_\beta{}^\gamma$ is a positive integer, or zero, and gives the number of ways in which the diagram $[\gamma]$ can be built on the diagram $[\beta]$ to yield the diagram $[\alpha]$. On the other hand $\{\alpha\}$ appears in $\Delta_{(1^l)}(L)$ with a frequency f_α, so that

1.7 $$f_\alpha = \sum_\beta f_\beta \sum_\gamma {}_\alpha\lambda_\beta{}^\gamma f_\gamma = \sum_\gamma f_\gamma \sum_\beta {}_\alpha\lambda_\gamma{}^\beta f_\beta.$$

It follows from 1.6 that

1.8 $$ {}_\alpha\lambda_\gamma{}^\beta = {}_\alpha\lambda_\beta{}^\gamma,$$

and we may write 1.7 in the form

1.9 $$f_\alpha = \sum_\beta \phi_\alpha{}^\beta f_\beta,$$

where

1.10 $$\phi_\alpha{}^\beta = \sum_\gamma {}_\alpha\lambda_\beta{}^\gamma f_\gamma.$$

If we denote the diagram conjugate to $[\alpha]$ by $[\alpha^*]$ then a little consideration will show that

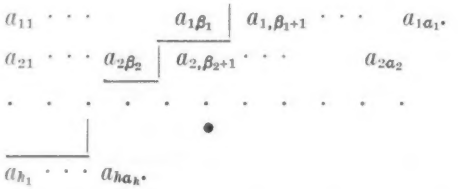
1.11 $$ {}_{\alpha^*}\lambda_{\gamma^*}{}^{\beta^*} = {}_\alpha\lambda_\gamma{}^\beta.$$

In the next section we shall see how the formulae 1.9 and 1.10 can be obtained from a different point of view.

2. In studying a representation ω of a finite group G it is fruitful to limit attention to a sub-group H of G and investigate the reducibility of ω with regard to H. Let us apply this procedure to S_l and, since our knowledge of the symmetric group is relatively so complete, let us consider the reducibility of a representation $[\alpha]$ of S_l with regard to a sub-group S_m, where $l = m + n$. In effect we restrict attention to permutations of S_l of the form

2.1 $$(1, 2, \cdots, m)(m+1, \cdots, l).$$

Such permutations form a sub-group of S_l which is the direct product of S_m and S_n. Let us suppose that the corresponding standard diagram $[\alpha]$ is of the form:

2.2 

From the restriction that the symbols must appear in the assigned or 'natural' order in any row or column of a standard diagram, it follows that the first m symbols must form a standard sub-diagram $[\beta]$ contained in $[\alpha]$. The remaining n symbols may be thought of as forming a standard *skew diagram,* which we may denote $[\alpha] - [\beta]$. It will be understood that this symbol has a meaning only if $[\beta]$ is contained in $[\alpha]$.

Young's rule for constructing the actual matrix representing the transposition $(a_r a_{r+1})$ in the orthogonal representation $[\alpha]$ of S_l is as follows:[5]

2.3 Put (i) unity in the leading diagonal where the associated standard diagram has a_r, a_{r+1} in the same row;

(ii) -1 in the leading diagonal where the associated standard diagram has a_r, a_{r+1} in the same column;

(iii) a quadratic matrix

$$\begin{pmatrix} -\rho & \sqrt{1-\rho^2} \\ \sqrt{1-\rho^2} & \rho \end{pmatrix},$$

[5] [5] reference Young [21], Part VI, pp. 217-8. Cf. also Thrall [6].

the elements being in the positions of intersections of a pair of rows and columns whose diagrams differ only by the exchange of a_r and a_{r+1}.

The quantity ρ is defined by the relation

$$\rho^{-1} = (\mu - \lambda) + (\gamma - \delta),$$

where $a_r = a_{\mu\delta}$ and $a_{r+1} = a_{\lambda\gamma}$, and we may suppose that $\mu > \lambda$ in the standard diagram in question. Young calls ρ the *axial projection* of $a_r a_{r+1}$ in $[\alpha]$.

It is important that ρ is defined in terms of the positions of the two symbols a_r, a_{r+1} *only* in $[\alpha]$. Consequently, the matrices representing the permutations of S_m are in reduced form *as they stand* and have as irreducible components all possible $[\beta]$ contained in $[\alpha]$. For exactly the same reason the matrices provide a representation of S_n, the variables corresponding to the ordinary standard diagrams being the standard skew diagrams. We can think of the formula 1.9 as representing the reduction of the direct product of S_m and S_n, where $\phi_\alpha{}^\beta$ is not only the frequency with which $[\beta]$ appears in $[\alpha]$ when S_l is restricted to S_m, but also the *degree* of a reducible representation of S_n, which may be denoted $[\alpha] - [\beta]$ and called a *skew* representation.

The skew diagram can be characterized as *standard* in just the same way as a right diagram, as already remarked. We have in fact a generalization which coincides with the right diagram under certain obvious conditions; e. g. in the case where $[\alpha] = [5, 4]$ and $[\beta] = [2^2]$.

3. The problem of determining the integers ${}_\alpha\lambda_\beta{}^\gamma$ is of considerable interest. What it amounts to is to establish a correlation between the two interpretations of the preceding paragraphs. We have on the one hand the $\phi_\alpha{}^\beta$ standard skew diagrams and we wish to associate them with the different direct products $\{\beta\} \times \{\gamma\}$ which yield contributions to the coefficients of $\{\alpha\}$ in 1.4. As in SG_1 § 5, MacMahon's lattice permutations [6] provide the clue.

Consider a permutation of the n letters

$$3.1 \qquad c_1{}^{\gamma_1} c_2{}^{\gamma_2} \cdots c_h{}^{\gamma_h},$$

where $\Sigma\gamma_i = n$. Such a permutation is said to be *lattice* if amongst the first r terms of it the number of c_1's $\geq$ the number of c_2's $\geq \cdots \geq$ the number of c_h's, for all r. If we add a second suffix to the c_i's according to the order

[6] Cf. also [2], pp. 67-71 and p. 95.

of their appearance for each i we may define the indices of the permutation. Considering first only the c_1's and the c_2's, if c_{2s} follows c_{1t} and precedes $c_{1,t+1}$ its index is defined to be

$$3.2 \qquad i_{12s} = s - t,$$

which may be positive, zero or negative. The resulting permutation of $c_1^{\gamma_1}c_2^{\gamma_2}$ is lattice if and only if no $i_{12s} > 0$. Similarly, we may define indices i_{23s}, i_{34s}, etc., and *a permutation of the letters* 3.1 *is lattice if and only if no* $i_{x,x+1,s} > 0$ *for* $x = 1, 2, \cdots, h-1$.

The important fact in this connection is that a non-lattice permutation may be made lattice according to a specific sequence of changes. This step-by-step process is as follows:

(a) Considering only the c_1's and the c_2's in the permutation, take the first c_2 with the greatest positive i_{12s} and change it into a c_1. Reallocating the second suffixes, repeat the process, continuing until all the c_1's and the c_2's are lattice.

(b) Considering only the c_2's and the c_3's in the permutation so modified, take the first c_3 with greatest positive i_{23s} and change it into a c_2. If this change upsets the 1-2 lattice property, correct for it by changing a c_2 into a c_1 according to (a); this may or may not be the new c_2. Reallocating the second suffixes, repeat the process, continuing until all the c_1's, c_2's and c_3's are lattice.

(c) Proceed as above until all the letters 3.1 are lattice.

It is obvious that any standard right diagram gives rise to a lattice permutation of the n letters 3.1 if we assume that the c_1's are the symbols appearing in the first row, the c_2's those in the second row, and so on. E. g. the diagram

$$\begin{matrix} 134 \\ 25 \end{matrix} \quad \text{gives rise to the permutation } c_1c_2c_1c_1c_2.$$

An exactly similar procedure is applicable to a standard skew diagram, but the resulting permutation will not necessarily be lattice. E. g. the skew diagram

$$\begin{matrix} & 2 \\ & 3 \\ 1 & \\ 4 & \end{matrix} \quad \text{gives rise to the permutation } c_3c_1c_2c_4.$$

Altogether, the skew diagram defined by $[\alpha] = [3^2, 2, 1]$ and $[\beta] = [2^2, 1]$

gives rise to $\phi_{\alpha}{}^{\beta} = 12$ permutations which are associated with the appropriate lattice permutation according to the following table.

Permutation	Lattice Permutation	Irreducible Representation
$c_1c_2c_3c_4$	$c_1c_2c_3c_4$	(1^4)
$c_1c_4c_3c_2$	$c_1c_2c_1c_1$	
$c_4c_1c_3c_2$	$c_1c_1c_2c_1$	$(3,1)$
$c_4c_3c_1c_2$	$c_1c_1c_1c_2$	
$c_3c_1c_4c_2$	$c_1c_1c_2c_2$	
$c_3c_4c_1c_2$	$c_1c_2c_1c_2$	(2^2)
$c_1c_2c_4c_3$	$c_1c_2c_3c_1$	
$c_1c_3c_4c_2$	$c_1c_2c_3c_1$	
$c_1c_3c_2c_4$	$c_1c_2c_1c_3$	
$c_1c_4c_2c_3$	$c_1c_2c_1c_3$	$(2,1^2)$ twice
$c_3c_1c_2c_4$	$c_1c_1c_2c_3$	
$c_4c_1c_2c_3$	$c_1c_1c_2c_3$	

We have in fact determined those irreducible representations $[\gamma]$ of S_n whose diagrams can, in the language of the Littlewood and Richardson theorem, be *built on* the diagram $[\beta]$ to yield $[\alpha]$, or can be *built in* $[\alpha] - [\beta]$. The proof given in SG_1 establishes the validity of this determination of the constants ${}_{\alpha}\lambda_{\beta}{}^{\gamma}$. In our example we have

$$\phi^{2^2,1}_{3^2,2,1} = f_{1^4} + f_{3,1} + f_{2^2} + 2f_{2,1^2}.$$

It should of course be pointed out that the same skew diagram may be defined by different pairs $[\alpha]$ and $[\beta]$. E. g.

$$\phi^{2^2,1}_{3^2,2,1} = \phi^{3^2,2,1}_{4^2,3,2}.$$

4. There is yet another approach to $\phi_{\alpha}{}^{\beta}$ which is of interest. Two formulae have long been known [7] to connect the f's of the symmetric group S_l with those of S_{l-1} and S_{l+1}:

4.1 $$f_{a_1,a_2,\ldots a_h} = f_{a_1-1,a_2,\ldots a_h} + f_{a_1,a_2-1,\ldots a_h} + \cdots + f_{a_1,a_2,\ldots a_h-1},$$

and

4.2 $$(l+1)f_{a_1,a_2,\ldots a_h} = f_{a_1+1,a_2,\ldots a_h} + f_{a_1,a_2+1,\ldots a_h} + \cdots + f_{a_1,a_2,\ldots a_h,1},$$

where the additions or deletions contribute only if they do not disturb the inequalities

[7] [5] reference Young [21], Part III, pp. 261-2.

$$\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_h.$$

Let us think of these relations as produced by the operation of adding or subtracting a node from the diagram $[\alpha]$. As such, the operation may be iterated to give

4.3 $$f_{\alpha(l)} = \sum_{\beta} \phi_{\alpha}{}^{\beta} f_{\beta(m)},$$

which is the same as 1.9, and

4.4 $$(l!/m!) f_{\beta(m)} = \sum_{\alpha} \phi_{\alpha}{}^{\beta} f_{\alpha(l)},$$

where $[\alpha(l)]$ contains $[\beta(m)]$ and $l = m + n$ as before. Iteration of either 4.3 or 4.4 requires that the ϕ's satisfy the relation

4.5 $$\phi_{\alpha(l)}{}^{\beta(m)} = \sum_{\mu} \phi_{\alpha(l)}{}^{\mu(r)} \cdot \phi_{\mu(r)}{}^{\beta(m)},$$

where $[\alpha(l)]$ contains $[\mu(r)]$ contains $[\beta(m)]$. Formula 4.4 is essentially numerical and has not the more general significance attributed to 4.3. Formula 4.5 is in fact a significant generalization of 4.3.

Young's proof of 4.1 is based on the enumeration of the standard diagrams. Since the last symbol a_l must occupy the end position in its row and column, after removing it, the remainder must still be standard. The inductive argument applied to the explicit form of f is somewhat involved, however, on account of the fact that one passes from the function $\Delta(x_1, x_2, \cdots, x_h)$ to $\Delta(x_1, x_2, \cdots, x_{h-1})$.

On the other hand the proof of 4.2 is much simpler. Noting that the expression for $f_{\beta(m)}$ is the coefficient of $x_1{}^{\beta_1+h-1} x_2{}^{\beta_2+h-2} \cdots x_h{}^{\beta_h}$ in the expansion of

$$(x_1 + x_2 + \cdots + x_h)^m \Delta(x_1, x_2, \cdots, x_h),$$

where Δ is the familiar alternating function, formula 4.2 is merely the coefficient of

4.6 $$x_1{}^{\beta_1+\overline{h-1}+1} x_2{}^{\beta_2+\overline{h-2}+1} \cdots x_h{}^{\beta_h+1} (x_1 + x_2 + \cdots + x_{h+1})$$

in the expansion of

4.7 $$(x_1 + x_2 + \cdots + x_{h+1})^{m+1} \Delta(x_1, x_2, \cdots, x_{h+1}).$$

By iteration, formula 4.4 is the coefficient of

4.8 $$x_1^{\beta_1+\overline{h-1}+n}x_2^{\beta_2+\overline{h-2}+n}\cdots x_h^{\beta_h+n}x_{h+1}^{n-1}\cdots x_{h+n-1}(x_1+x_2+\cdots+x_{h+n})^n$$

in the expansion of

4.9 $$(x_1+x_2+\cdots+x_{h+n})^l\Delta(x_1,x_2,\cdots,x_{h+n}).$$

The extra terms which appear in 4.6 have zero coefficient in 4.7. A similar argument is applicable to 4.8 if the factors $(x_1+x_2+\cdots+x_{h+n})$ are thought of as multiplied in succession.[8]

One would like to have an explicit expression for ϕ analogous to that for f. If none of the appropriate terms in 4.8 had vanishing coefficients in 4.9, then ϕ would be a multinomial coefficient. In general, however, this is not the case, so that

4.10 $$\phi_{\alpha(l)}{}^{\beta(m)} \leq n!/\Pi(\alpha_i-\beta_i)!,$$
$$\text{and also} \leq n!/\Pi(\alpha^*_i-\beta^*_i)!,$$

from 1.11.

It is worth pointing out that a skew diagram need *not* be *connected*, in the sense that any one of its nodes can be reached from any other by a succession of horizontal and vertical steps. Clearly a right diagram is connected. Let us suppose that $[\alpha]-[\beta]$ breaks up into a number of connected skew diagrams containing respectively $p, q, r, \cdots$ nodes, so that $n=p+q+r+\cdots$. Then a little consideration will show that

4.11 $$\phi_{\alpha(l)}{}^{\beta(m)} = (n!/p!q!r!\cdots)\phi_{(p)}\cdot\phi_{(q)}\cdot\phi_{(r)}\cdots,$$

where the multinomial coefficient comes from the division of the n symbols into sets of $p, q, r, \cdots$ symbols. After such division, the symbols in any set can be taken in their natural order and arranged in the corresponding skew diagram in just ϕ distinct ways.

In concluding this discussion of the generation of the skew representations it should be pointed out that the idea of representing the addition or subtraction of a node from a diagram as an operator goes back to Young (Q. S. A. IV, §§ 6-11) who used the notation E_r and Δ_r for the operation as applied to the r-th row of a diagram. Later (Q. S. A. VI, § 8) he extended the idea, considering the operator $\Delta_i\Delta_j$ applied to a substitutional expression T. In the Littlewood and Richardson Theorem we have the first significant application of the method. One might say that Nakayama's introduction of the notion of a *hook*, to be discussed in Part II, is the second.

[8] Cf. [2], p. 68.

PART II.

The Murnaghan-Nakayama Recursion Formula.

5. We now apply these ideas to the concept of a *hook* as introduced by Nakayama.[4] Consider a diagram $[\alpha]$. A *hook* H_r in $[\alpha]$ is defined in § 1 of N_1 to be any diagram

5.1 $$H_r = [n-r, 1^r], \qquad (r = 0, 1, \cdots n-1),$$

whose horizontal (vertical) nodes are the *last* $n-r(r)$ nodes of some row (the appropriate column) of $[\alpha]$. We shall say that this hook is *even* or *odd* according as r is even or odd. It will sometimes be convenient to speak of the horizontal part of H_r of length $n-r$ as the *arm* and of the vertical part of length r as the *leg*. As an example consider the diagram $[\alpha] = [4^2, 3, 2, 1]$:

5.2

. . . o
. . o o
. o o
. o
.

There is just one hook of length 6, namely $[3, 1^3]$. If this hook is *removed* from $[\alpha]$, the procedure is to think of those nodes which are isolated as being moved to the left and up to produce the diagram $[3, 2, 1^3]$. On the other hand Nakayama remarks in § 5 of N_1 that it is sometimes necessary to think of removing the *equivalent 'rim'* of the diagram, as indicated in 5.2. This latter interpretation is more significant for the operational approach developed in Part I of the present paper. We shall speak of the rim of the diagram as a *skew hook*, to distinguish it from its equivalent *right hook* (5.1). Clearly, the *skew* hook in 5.2 is precisely the skew diagram $[4^2, 3, 2, 1] - [3, 2, 1^3]$.

Nakayama's purpose in introducing the notion of a hook was to study the modular properties of the characters of the symmetric group. We shall postpone the application of our methods to such modular considerations and consider here only the application of the theory to the proof of the Murnaghan-Nakayama recursion formula, as given in § 9 of N_1:

Let $H_1, H_2, \cdots$ be the totality of hooks of length n in the diagram $T = [\alpha]$, then

5.3 $$\chi(T : P) = \sum (-1)^{r_i} \chi(T - H_i : P'),$$

where P is any permutation of S_l which contains a cycle C of length n and P' is the permutation on m symbols obtained by removing C from P.

Clearly, the permutation P is contained in the direct product of S_m and S_n and its character in $[\alpha]$ is given by the analogue of 1.9 or 4.3 for characters:

5.4 $$\chi(T;P) = \chi_\alpha(P) = \sum_\beta (\chi(C) \text{ in } [\alpha] - [\beta] \cdot \chi_\beta(P').$$

6. We have from **1** that

1.10 $$[\alpha] - [\beta] = \sum_\gamma {}_\alpha\lambda_\beta{}^\gamma [\gamma].$$

Murnaghan [9] proved that

6.1 $$\chi_{\gamma(n)}(C) = (-1)^r \text{ or } 0,$$

according as $[\gamma]$ is an H_r or not. The following theorem gives the requisite information concerning the ${}_\alpha\lambda_\beta{}^\gamma$ in 1.10.

THEOREM. *If the skew-diagram $[\alpha] - [\beta]$:*

(i) *contains interior nodes of $[\alpha]$, then no hook representation occurs;*

(ii) *is a complete skew hook, then there is just one hook representation;*

(iii) *is an incomplete hook, then the number of even hook representations is equal to the number of odd ones.*

If a node of $[\alpha]$ is not on the rim of the diagram it may be said to be an *interior* node of $[\alpha]$.

If $[\alpha] - [\beta]$ contains an interior node of $[\alpha]$, then it must contain the configuration

6.2 $$\begin{matrix} \circ & \cdot \\ \cdot & \cdot \end{matrix},$$

where the interior node is ringed. A little consideration will show that no four elements of H_r can be arranged in the pattern 6.2 without violating 1.5, which proves (i). As an illustration we have:

6.3 $$\phi^{2^3,1^2}_{4^2,3,2,1} = f_{3,2,1} + f_{2^3} + f_{2^2,1^2} = 30.$$

[9] [3], p. 462.

If $[\alpha] - [\beta]$ is a skew hook, then certainly the equivalent H_r will appear as an irreducible component of $[\alpha] - [\beta]$; since the number of horizontal and vertical steps is just accounted for, the corresponding ${}_\alpha\lambda_\beta{}^\gamma = 1$, and no other hooks can be built in the skew diagram without violating 1. 5. This proves (ii). E. g.:

$$6.4 \qquad \phi^{3,2,1^3}_{4^2,3,2,1} = f_{3,1^3} + 2f_{3,2,1} + f_{3^2} + f_{2^3} + f_{2^2,1^2} = 61,$$

where $H_r = [3, 1^3]$ as in 5. 2.

Case (iii) is a little more complicated. If one thinks of the removal of a node from a skew hook $[\alpha] - [\beta]$, two possibilities arise: (a) the node in question may be at one end, or (b) it may be an internal node of the rim of $[\alpha]$. Possibility (a) leaves the hook still a hook on one fewer symbols and so is of no interest here. Possibility (b) relaxes two conditions, since the removal must be at a corner of the skew diagram. Two hooks H_{r_1} and H_{r_2} may be built in the resulting skew diagram, which we may call an *incomplete* skew hook, one obtained by removing a node from the arm of H_r, the other by removing a node from its leg. As an illustration of the process we have:

$$6.5 \qquad \phi^{4,3,1}_{4^2,3,2,1} = f_{4,2} + f_{4,1^2} + 3f_{3,2,1} + f_{3^2} + f_{3,1^3} + f_{2^3} + f_{2^2,1^2} = 96,$$

where $H_r = [4, 1^3]$, $H_{r_1} = [4, 1^2]$ and $H_{r_2} = [3, 1^3]$. One of H_{r_1} and H_{r_2} is consequently *even* while the other is *odd*.

If now a second node be removed from this incomplete skew hook, it may be next to the first node or it may be another internal node of the rim of $[\alpha]$. In the first case the only possible hooks which can be built in $[\alpha] - [\beta]$ are obtained by removing a node from the arms of *both* H_{r_1} and H_{r_2} or from their legs; one of the resulting hooks will be even while the other one will be odd. The removal of a second internal node again relaxes two conditions and makes it possible to build hooks obtained by removing a node from *each* of the arms and legs of H_{r_1} and H_{r_2}; of the resulting four hooks, two will be even and two odd. E. g. we have:

$$6.6. \qquad \phi^{3^2,2}_{4^2,3,2,1} = f_{4,2} + f_{4,1^3} + f_{3^2} + 3f_{3,2,1} + 2f_{3,1^3} + f_{2^3} + 2f_{2^2,1^2} + f_{2,1^4} = 120,$$

where $[4, 1^2]$, $[3, 1^3]$ and $[3, 1^3]$, $[2, 1^4]$ are obtained as above described from $[4, 1^3]$ and $[3, 1^4]$. Further repetition of the argument proves (iii).

7. The Murnaghan-Nakayama recursion formula now follows immediately from 6. 1, since

7.1 $$\chi(C) \text{ in } [\alpha] - [\beta] = 0$$

in cases (i) and (iii), while

7.2 $$\chi(C) \text{ in } [\alpha] - [\beta] = (-1)^r$$

in case (ii).

Taking $n = p$, it follows from the modular theory that

7.3 $$f_{\gamma(p)} \equiv (-1)^r \text{ or } 0, \text{ mod } p,$$

according as $[\gamma]$ is an H_r or not. Formula 7.3 should be compared with 6.1. As before

7.4 $$\phi_{\alpha(l)}{}^{\beta(m)} = \sum_{\gamma} {}_{\alpha}\lambda_{\beta}{}^{\gamma} f_{\gamma(p)} \equiv (-1)^r \text{ or } 0, \text{ mod } p,$$

according as $[\alpha] - [\beta]$ is a skew hook equivalent to H_r or not. The analogy between the f's and the ϕ's is striking throughout.

THE UNIVERSITY OF TORONTO.

REFERENCES.

[1] Stig Comét, "Une propriété des déterminantes et son application au calcul des caractères des groupes symétriques," *Kungl. Fysiografiska Sällskapets i Lund Förhandlingar* 14, no. 7, pp. 84-94 (1945).

[2] D. E. Littlewood, *The Theory of Group Characters*, Oxford, 1940.

[3] F. D. Murnaghan, "On the representations of the symmetric group," *American Journal of Mathematics*, vol. 59 (1937), pp. 437-488.

[4] T. Nakayama, "On some modular properties of irreducible representations of a symmetric group," Part I, *Japanese Journal of Mathematics*, vol. 17 (1940), pp. 89-108. Part II, same volume, pp. 411-423.

[5] G. de B. Robinson, "On the representations of the symmetric group," *American Journal of Mathematics*, vol. 60 (1938), pp. 745-760.

[6] R. M. Thrall, "Young's semi-normal representation of the symmetric group," *Duke Mathematical Journal*, vol. 8 (1941), pp. 611-624.

[7] H. Weyl, *The Classical Groups*, Princeton, 1939.

POLYNOMIAL MATRICES IN ONE VARIABLE, DIFFERENTIAL EQUATIONS AND MODULE THEORY.*

By Ernst Snapper.

Introduction and Summary. Let $P[x]$ be the ring of polynomials in one variable x with coefficients in the commutative field P. Let A be an $m \times s$ matrix of rank r whose elements belong to $P[x]$. Let Δ_i be the highest common factor of the $i \times i$ subdeterminants of A for $i = 1, \cdots, r$. The purpose of Part I of this paper is to develop the theory of the norm $\Delta = \Delta_r$ and the elementary divisor $\epsilon = \Delta_r : \Delta_{r-1}$ intrinsically in terms of the column space M of A (the column space of a matrix is the vector space generated by its columns). The purpose of Part II is to extend this theory to the case where M is any module with a Noetherian ring (a ring in which every ideal has a finite basis) as operator domain.

In Part I it is shown how Δ and ϵ can be defined, by the use of composition sequences and other special sequences, intrinsically in terms of M, thus making it obvious that two matrices with the same column spaces have the same Δ and ϵ. These definitions are then used to investigate the properties of Δ and ϵ (see **1.1**) and to develop the theory of a system of linear algebraic equations with coefficients in $P[x]$ (see **1.2** and **1.3**) and of a system of homogeneous linear differential equations with constant coefficients (see **1.4-1.6**). This treatment of the matrix A differs from the classical theory in the following respects:

1) In the classical theory Δ and ϵ are defined in terms of determinants of A and not intrinsically in terms of M (see [1], p. 91 or [2], p. 27; square brackets refer to the references).

2) In the classical theory of the above mentioned differential equations, all the determinantal factors $\Delta_r, \Delta_{r-1}, \cdots, \Delta_1$ or elementary divisors $\Delta_r : \Delta_{r-1}$, $\Delta_{r-1} : \Delta_{r-2}, \cdots, \Delta_1$ occur (see [1], chapter 5) while the matrix of the system is assumed to be square and non-singular. In **1.4-1.6** this theory is derived by the use of only Δ and ϵ while composition sequences of M occur instead of determinantal factors or elementary divisors of order less than r. Furthermore, in **1.4** the notion of "trivial solution" is introduced which enables us to develop the theory for a system with an arbitrary $m \times s$ matrix instead of

* Received October 4, 1945.

only for a system with a square non-singular matrix. In these three sections it is assumed that P is the field of complex numbers, not from mathematical necessity but in the hope that these sections will also be read by non-algebraists who are interested in differential equations.

In Part II it is shown how, by the use of the module theory developed by P. M. Grundy (see [5]) and the author (see [6]), the theory of Part I can be extended to the case where M is any module with a Noetherian ring $\mathfrak{s}$ as operator domain. Sections **2. 1-2. 4** review those parts of this module theory, necessary for the extensions. Section **2. 5** shows how the notions of Δ and ϵ can be extended. Sections **2. 6-2. 8** deal with the theory of a system of algebraic linear equations

$$(1) \qquad \Sigma^{s}_{j=1}\alpha_{ij}x_j = \gamma_i, \qquad i = 1, \cdots, m$$

where α_{ij} and γ_i are given elements of $\mathfrak{s}$. The "criterion of solvability" which says when (1) can be solved by elements $x_j \in \mathfrak{s}$ is based on the "criterion of lengths" (see **2. 6**) which is the module analogue of the "Langensatz" of ideal theory (see [7], p. 35). It is shown in **2. 7** how the "criterion of lengths" becomes the classical theorem of solvability of the system (1) in the cases where $\mathfrak{s}$ is specialized to be a field (see [2], p. 70) or a principal ideal ring or $P[x]$ (see [3], p. 60) or the integers of an algebraic number field (see [4], p. 340).

A paper to be published in this journal under the name "Polynomial Matrices in several Variables," referred to here as MSV, will treat the special case where $\mathfrak{s} = P[x_1, \cdots, x_n]$, that is a polynomial ring in several variables. Frequent remarks throughout the present paper point out connections between Part I, Part II, MSV, [5] and [6]; they are not a logical part of the development of the theory.

The author is not able to extend the notions of the determinantal factors $\Delta_{r-1}, \cdots, \Delta_1$ and elementary divisors $\Delta_{r-1} : \Delta_{r-2}, \cdots, \Delta_1$ of order *less* than r in an analogous way beyond the case where $\mathfrak{s} = P[x]$. Such an extension would be useful only if it supplied a criterion by which one could tell when two modules are operator isomorphic; such a criterion is indeed supplied by the elementary divisors when $\mathfrak{s} = P[x]$.

Part I. Polynomial Matrices in One Variable.

1. 1. Closure, associated primes, elementary divisor and norm of a module. Let $P[x]$ be the polynomial ring consisting of the polynomials in one variable x whose coefficients are elements of the arbitrary, commutative,

field P. Elements of $P[x]$ are called *scalars* and elements of P *constants.* Let V be the m-dimensional vector space V over $P[x]$, consisting of the column vectors whose m components are elements of the scalar domain $P[x]$. The vectors of V can be added, subtracted and multiplied by scalars in the usual way. A *module* is a subset of V which is closed under vector subtraction and scalar multiplication. If, without further explanation, a capital Roman letter, a lower case Roman letter or a lower case Greek letter is used, the letter always denotes, respectively, a module, a vector or a scalar. The ideal of $P[x]$, generated by α, is denoted by (α) and the module, generated by $u, v, \cdots, w$, is denoted by $(u, v, \cdots, w)$. In the same way $(M, u, v, \cdots, w)$ denotes the module which is generated by the vectors of M together with $u, v, \cdots, w$; i. e., $h \in (M, u, v, \cdots, w)$ if $h = m + \alpha u + \beta v + \cdots + \gamma w$ where $m \in M$. We shall never distinguish between two scalars ϕ_1 and ϕ_2 if $\phi_1 = \alpha\phi_2$ where α is a constant. Hence, we shall speak about "the" generator of an ideal of $P[x]$ and shall never worry about the fact that such a generator is determined only up to a multiplicative constant. The symbol $M : \alpha$ denotes the module consisting of all vectors v such that $\alpha v \in M$. Clearly, $M : \alpha \supseteq M$. The symbol $M : N$ denotes the ideal consisting of all scalars α such that $\alpha n \in M$ for all $n \in N$. *Throughout this paper, the symbols* $\subset$ *and* $\supset$ *will be used exclusively for proper inclusion.*

DEFINITION **1.11.** *An irreducible scalar* $\pi \neq 0$ *is a non-zero associated prime of* M *if and only if* $M : \pi \supset M$.

Remark **1.11.** The non-zero associated primes of Definition **1.11** are identical with those defined by the Noether decomposition of M (see Remark **2.23**). For the zero associated prime of a module, see Remark **2.32**.

DEFINITION **1.12.** *The closure* Cl (M) *is the module which consists of all vectors* v *such that* $\alpha v \in M$ *for some non-zero scalar. The elementary divisor* ϵ *of* M *is the generator of the ideal* M : Cl (M).

Clearly $\epsilon \neq 0$, since if Cl $(M) = (v_1, \cdots, v_s)$ there exist non-zero scalars $\alpha_1, \cdots, \alpha_s$ such that $\alpha_i v_i \in M$ for $i = 1, \cdots, s$. The product $\alpha_1 \cdot \cdots \cdot \alpha_s$ is then a non-zero element of M : Cl (M), hence $\epsilon \neq 0$. It follows immediately that $M : \epsilon =$ Cl (M). Furthermore always $M \subseteq$ Cl (M). We say that M is *closed* if $M =$ Cl (M) and that M is *dense* if Cl $(M) = V$.

Remark **1.12.** If A is a matrix whose columns generate M, then ϵ of M is equal to the usual elementary divisor of highest rank of A (see [3], Theorem II).

The following theorem shows that the number of non-zero associated primes of M is finite.

THEOREM 1.11. *If M is not closed, the non-zero associated primes are the irreducible factors of ϵ. Furthermore, the following three statements are equivalent:* (1) *M is closed;* (2) *ϵ is a non-zero constant;* (3) *M has no non-zero associated primes.*

Proof. If π is a non-zero associated prime of M, a $v \notin M$ exists such that $\pi v \in M$ and $\epsilon v \in M$. If π were not a factor of ϵ, we could find scalars α and β such that $(\alpha\pi + \beta\epsilon)v = v$ which would imply $v \in M$. Hence, every non-zero associated prime is a factor of ϵ. Conversely, let π be an irreducible factor of ϵ; hence $\epsilon = \pi^\rho\phi$ and, if $v \in \mathrm{Cl}\,(M)$, $\pi^\rho\phi v \in M$. Now, if $M : \pi = M$, then we could conclude from $\pi^\rho\phi v \in M$ that $\phi v \in M$. Since ϵ is the highest common factor of all scalars which transform all vectors of Cl (M) into vectors of M, this would imply that $\phi = \epsilon$ against hypothesis. Consequently, $M : \pi \supset M$ and π is a non-zero associated prime of M. For the second part of the theorem, suppose that M is closed. Then $M : \mathrm{Cl}\,(M) = M : M = P[x]$ and hence ϵ is a non-zero constant. Conversely, if ϵ is a non-zero constant and $v \in \mathrm{Cl}\,(M)$, then from $\epsilon v \in M$ follows $v \in M$ and hence Cl $(M) = M$. Consequently, (1) and (2) are equivalent. We now show that (1) and (3) are equivalent. If M is closed, $M : \alpha = M$ for any non-zero scalar α and hence (3) is satisfied. Conversely, if (3) is satisfied, M must be closed for, if $M \subset \mathrm{Cl}\,(M)$, then $M : \mathrm{Cl}\,(M) \subset P[x]$ and hence $\epsilon \notin P$. Every irreducible factor of ϵ is then a non-zero associated prime of M according to the first part of the theorem, against the hypothesis that (3) holds. This proves Theorem 1.11.

Let $\epsilon = \Pi^h_{j=1}\pi_j^{\rho_j}$. The meaning of the irreducible factors π_j has been explained. The following theorem explains the meaning of the multiplicity ρ_j.

THEOREM 1.12. *The sequence* (1) $M \subset M : \pi_j \subset \cdots \subset M : \pi_j^{\rho_j}$ *consists of $\rho_j + 1$ different terms. For $h \geq 1$, $M : \pi_j^{\rho_j+h} = M : \pi_j^{\rho_j}$.*

Proof. In order to show that the $\rho_j + 1$ terms of (1) are different, we observe that $M : \pi_j^{k+1} = (M : \pi_j^k) : \pi_j$ and hence that all we have to show is $M : \pi_j^{\rho_j-1} \subset M : \pi_j^{\rho_j}$. Since $\epsilon = \pi_j^{\rho_j}\phi$, we can find a $v \in \mathrm{Cl}\,(M)$ such that $\pi_j^{\rho_j-1}\phi v \notin M$ while of course $\pi_j^{\rho_j}\phi v \in M$, for, otherwise, $\epsilon = \pi_j^{\rho_j-1}\phi$. Hence, $\phi v \in M : \pi_j^{\rho_j}$ but $\phi v \notin M : \pi_j^{\rho_j-1}$ which proves $M : \pi_j^{\rho_j-1} \subset M : \pi_j^{\rho_j}$. Now, let $v \in M : \pi_j^{\rho_j+h}$ where $h \geq 1$. Then, $\pi_j^{\rho_j+h}v \in M$ and $\pi_j^{\rho_j}\phi v \in M$. Since π_j and ϕ are relatively prime we can find scalars α and β such that $(\alpha\pi_j^h + \beta\phi)v = v$ which implies $\pi_j^{\rho_j}v = \alpha\pi_j^{\rho_j+h}v + \beta\pi_j^{\rho_j}\phi v \in M$ and hence $v \in M : \pi_j^{\rho_j}$.

We now turn our attention to the norm Δ of M by refining the sequence (1) to a composition sequence. Here, a sequence $M_0 = M \subset M_1 \subset \cdots \subset M_l$ is called a *composition sequence from M to M_l of length l* if no module N exists such $M_i \subset N \subset M_{i+1}$ for some $0 \leq i \leq l-1$. Clearly, the factor modules $0 \subset M_1/M \subset \cdots \subset M_l/M$ form an ordinary composition sequence of length l of M_l/M (see [9], p. 150). Conversely, if $0 \subset C_1 \subset \cdots \subset M_l/M$ is a composition sequence of length l of M_l/M and if M_i consists of all the vectors of the coset C_i, then $M \subset M_1 \subset \cdots \subset M_l$ is a composition sequence from M to M_l of length l.

THEOREM 1.13. *If $M \subset N \subseteq \mathrm{Cl}(M)$ and if $M \subset M' \subset \cdots \subset N$ is any sequence from M to N, this sequence can be refined to a composition sequence from M to N. The length l of such a composition sequence depends only on M and N and, if $M_i \subset M_{i+1}$ are two consecutive members of the sequence, the factor module M_{i+1}/M_i has finite rank r_i with respect to P. Consequently, N/M has finite rank $r = \Sigma^l_{i=1} r_i$ with respect to P.*

Proof. For "rank with respect to P" we shall write P-rank. Theorem 1.13 follows from the theory of composition sequences of $\mathrm{Cl}(M)/M$ as soon as we have proved the existence of one composition sequence from M to $\mathrm{Cl}(M)$, of which two consecutive members give rise to a factor module which has finite P-rank. Let $\epsilon = \Pi^h_{j=1} \pi_j^{\rho_j}$ and consider the sequence $M \subset M : \pi_1^{\rho_1} \subset M : \pi_1^{\rho_1} \pi_2^{\rho_2} \subset \cdots \subset M : \epsilon = \mathrm{Cl}(M)$. If $M_i = M : \pi_1^{\rho_1} \cdots \pi^{\rho_i}$, then $M_{i+1} = M_i : \pi_{i+1}^{\rho_{i+1}}$ for $0 \leq i \leq h-1$ and the sequence $M_i \subset M_{i+1}$, formed by two consecutive members, can be refined to $M_i \subset M_i : \pi_{i+1} \subset M_i : \pi^2_{i+1} \subset \cdots \subset M_i : \pi_{i+1}^{\rho_{i+1}} = M_{i+1}$. We show that we can draw a composition sequence from $M_i : \pi^k{}_{i+1}$ to $M_i : \pi^{k+1}_{i+1}$ for $0 \leq k \leq \rho_{i+1} - 1$. Suppose that $v \in M_i : \pi^{k+1}_{i+1}$ and $v \notin M_i : \pi^k{}_{i+1}$ and let $L = ((M_i : \pi^k{}_{i+1}), v)$. Since $\alpha v \in M_i : \pi^k{}_{i+1}$ if and only if $\alpha \in (\pi_{i+1})$, the factor module $L/M_i : \pi^k{}_{i+1}$ is operator isomorphic with $P[x]/(\pi_{i+1})$ with respect to $P[x]$ as operator domain. Since (π_{i+1}) is a maximal prime ideal of $P[x]$, $P[x]/(\pi_{i+1})$ has no proper non-zero sub-ideal and has finite P-rank ∂_{i+1} if ∂_{i+1} is the degree of the polynomial π_{i+1}. Hence, no proper sub-module can be included between M and L and ∂_{i+1} is the P-rank of L/M. If $L = M_i : \pi^{k+1}_{i+1}$, the proof is complete; otherwise we choose a vector $w \in M_i : \pi^{k+1}_{i+1}$ and $w \notin L$ and repeat the process. Because the ascending chain condition holds in V, we must reach $M_i : \pi^{k+1}_{i+1}$ in a finite number of steps and Theorem 1.13 is proved.

Since $M : \pi_j^{\rho_j} \subseteq \mathrm{Cl}(M)$, we can draw composition sequences from M to $M : \pi_j^{\rho_j}$ according to Theorem 1.13. In the remainder of Part I, l_j will

always denote the length of a composition sequence from M to $M : \pi_j^{\rho_j}$ and ∂_j the degree of the polynomial π_j for $1 \leq j \leq h$.

THEOREM 1.14. *The P-rank of* $M : \pi_j^{\rho_j}/M$ *is* $l_j\partial_j$ *and of* Cl $(M)/M$ *is* $\partial = \Sigma^h_{j=1} l_j\partial_j$.

Proof. According to the notation of the previous theorem, two consecutive members $M_i \subset M_{i+1}$ of the sequence $M \subset M_1 \subset \cdots \subset M_h = \text{Cl}\,(M)$ were refined to a composition sequence of length l'_{i+1}, and $l'_{i+1}\partial_{i+1}$ was the P-rank of M_{i+1}/M_i. Hence, all we have to show is that $l'_{i+1} = l_{i+1}$. Let $M'_j \subset M'_{j+1}$ be two consecutive members of a composition sequence from M to $M : \pi_{i+1}^{\rho_{i+1}}$ and let $\pi_1^{\rho_1} \cdots \pi_i^{\rho_i} = \phi_i$. Then, $M'_{j+1} = (M'_j, v_j)$ and $\alpha v_j \in M'_j$ if and only if $\alpha \in (\pi_{i+1})$. Since ϕ_i and π_{i+1} are relatively prime, we can find scalars β and γ such that $(\beta\phi_i + \gamma\pi_{i+1})v_j = v_j$. Hence $M'_{j+1} = (M'_j, \phi_i w_j)$ where $w_j = \beta v_j$. The composition sequence from M to $M : \pi_{i+1}^{\rho_{i+1}}$ can hence be written as $M \subset (M, \phi_i w_1) \subset (M, \phi_i w_1, \phi_i w_2) \subset \cdots \subset (M, \phi_i w_1, \cdots, \phi_i w_l) = M : \pi_{i+1}^{\rho_{i+1}}$. It can then be proved without difficulty that $M'_i \subset (M_i, w_1) \subset (M_i, w_1, w_2) \subset \cdots \subset (M_i, w_1, \cdots, w_l) = M_i : \pi_{i+1}^{\rho_{i+1}}$ is a composition sequence from M_i to $M_i : \pi_{i+1}^{\rho_{i+1}}$ and hence that $l'_{i+1} = l_{i+1}$.

DEFINITION 1.13. *The norm* Δ *of a module* M *is defined as* $\Delta = \Pi^h_{j=1}\pi_j^{l_j}$.

Since ρ_j is the length of an ordinary sequence (without repetition) from M to $M : \pi_j^{\rho_j}$ and l_j is the length of a composition sequence, $l_j \geq \rho_j$ and hence ϵ is always a divisor of Δ. According to Theorem 1.14, the degree of Δ is ∂, that is the P-rank of Cl $(M)/M$.

Remark 1.13. If A is a matrix whose columns generate M, Δ of M is equal to the determinantal factor of highest rank of A as follows easily from [3], pp. 60 and 61. The difference in character between ϵ and Δ is now clear. The multiplicities of Δ are "lengths" and hence are closely related to P-ranks, which makes Δ fundamental for systems of algebraic equations. The multiplicities of ϵ, however, tell when certain sequences break off, which makes ϵ fundamental for systems of differential equations.

Remark 1.14. The composition sequences from M to $M : \pi_j^{\rho_j}$ correspond to the primary composition sequences of the general module theory (see 2.4 and 2.5) which, in their turn, correspond to the primary composition sequences of ideal theory (see [10], Section 23).

1.2. Rank and closure. Let M be a module, Cl (M) its closure, and r its rank; i.e., r is the maximal number of vectors of M which are linearly

independent with respect to $P[x]$. Hence, rank, without further explanation, always means $P[x]$-rank. The following theorem connects the notions of closure and rank.

THEOREM 1.21. *Cl (M) is closed, has the same rank r as M, and contains every module whose rank is r and which contains M. Furthermore,* Cl $(M_1 \cap M_2) =$ Cl $(M_1) \cap$ Cl (M_2) *and, if* $M_1 \subseteq M_2$, *then* Cl $(M_1) \subseteq$ Cl (M_2).

Proof. The first half of the theorem is proved in [3], p. 56 and the second half in [6], p. 259.

It follows from Theorem 1.21 that Cl (M) could have been defined as the only closed module of rank r which contains M. Consequently, if $M_1 \subseteq M_2$ and M_1 and M_2 have the same rank, then Cl $(M_1) =$ Cl (M_2). Finally, M is dense if and only if $r = m$.

Remark 1.21. Theorem 1.21 is used in **1.3** for the theory of systems of linear algebraic equations. In the case of arbitrary modules the notion of rank is replaced by the more general notion of length of a primary sequence (see **2.7** and **2.8**).

DEFINITION 1.21. *If A is an $m \times s$ matrix whose columns are vectors of V, then the closure, rank, elementary divisor, norm and associated primes of A are defined as those of the column space of A, i.e., of the module M generated by the columns of A.*

1.3. Systems of linear equations.

THEOREM 1.31. *Let $M_1 \subseteq M_2$ be two modules. Then, $M_1 = M_2$ if and only if M_1 and M_2 have the same rank and norm.*

Proof. It is obvious that if $M_1 = M_2$ the norms and ranks of M_1 and M_2 are the same. If $M_1 \subseteq M_2$ and the ranks are equal, then Cl $(M_1) =$ Cl (M_2). Let $M_2 \subset M_{21} \subset \cdots \subset M_{2l} =$ Cl (M_2) be a composition sequence from M_2 to Cl (M_2) of length l and let r_i be the P-rank of the factor module $M_{2,i+1}/M_{2,i}$. Then, $\partial = \sum_{i=1}^{l} r_i$ is the P-rank of Cl $(M_2)/M_2$ and hence is equal to the degree of the norm of M_2. If $M_1 \subset M_2$, the sequence $M_1 \subset M_2 \subset M_{21} \subset \cdots \subset M_{2l} =$ Cl (M_1) would prove that the P-rank of Cl $(M_1)/M_1$ is $\partial + r_0$ where r_0 is the P-rank of M_2/M_1. Hence, if also the norms of M_1 and M_2 are the same, $r_0 = 0$ and $M_1 = M_2$.

Theorem 1.31 contains the criterion, given in Theorem 1.32, for the

solvability of a system of linear equations (Theorem **1.32** is the analogue of the theorem on p. 340 in [4]).

THEOREM **1.32**. *Let* $\Sigma_{j=1}^{s}\alpha_{ij}z_j = \gamma_i$, *where* $i = 1, \cdots, m$ *and* α_{ij} *and* γ_i *are polynomials of* $P[x]$, *be a system of* m *linear equations for the* s *unknowns* z_j. *Then, this system has a solution* $z_j \in P[x]$ $(j = 1, \cdots, s)$ *if and only if the* $m \times s$ *matrix* $A = (\alpha_{ij})$ *and the augmented* $m \times (s+1)$ *matrix* $B = (\alpha_{ij}, \gamma_i)$ *have the same rank and norm.*

Proof. Let M be the column space of A and g the vector whose components are $\gamma_1, \cdots, \gamma_m$. Then the system has a solution if and only if $g \in M$, i. e., if and only if $M = L$, where $L = (M, g)$. Since $M \subseteq L$, Theorem **1.32** is an immediate corollary of Theorem **1.31** and Definition **1.21**.

1.4. Systems of differential equations. We shall investigate the exponential solutions of a system S of linear, homogeneous, differential equations with constant coefficients.

Let W be the m-dimensional vector space consisting of the column vectors whose components are elements of the polynomial ring $P[t]$, where t is a new variable. For reasons given in the introduction, we assume in **1.4**, **1.5** and **1.6** that P *is the field of complex numbers.* An exponential vector $w \exp(\xi t)$ is the product of a vector $w \in W$ and an exponential $\exp(\xi t) = e^{\xi t}$, where $\xi \in P$ and where e is the basis of the natural logarithms. The derivative with respect to t, $Dw \exp(\xi t)$, is defined as in analysis where it is proved that $Dw \exp(\xi t) = \exp(\xi t)\ (D + \xi)w$ (see [1], Chapter V). If it is important to indicate the independent variable, we will write $A(x)$ instead of A and $w(t)$ instead of $w \in W$. For $v \in V$ we may write $v(x)$, and $v(t)$ then arises from $v(x)$ by the substitution $x = t$. In the same way, the operator vector $v(D)$ is the vector whose components are differential operators which arise from $v(x)$ or $v(t)$ by the substitutions $x = D$ and $t = D$ respectively. We assume in **1.4**, **1.5** and **1.6** that A *is a fixed* $m \times s$ *matrix of rank* r whose columns are vectors of V. The operator matrix $A(D)$ arises from A by the substitution $x = D$. The system S whose exponential solutions we want to investigate is given by $A(D)y = 0$. Hence we want to discuss the exponential vectors $y = w \exp(\xi t)$ such that $A(D)w \exp(\xi t) = 0$.

An exponential vector $w \exp(\xi t)$ is said to be of degree κ if κ is the degree of w, where the degree of a vector is the highest degree of the components of the vector. Hence, a vector $w(t)$ is of degree κ if $w = \Sigma_{j=0}^{\kappa} c_j t^j$ where the c_j's are constant vectors (i. e., the components of c_j are elements of P) and where $c_\kappa \neq 0$. The letter c (or c_j etc.), used without further

explanations, will always denote a constant vector or a constant, depending on the connotation.

LEMMA 1.41. *Every exponential vector* $w\exp(\xi t)$ *of degree* κ *can be written as* $u(D)t^\kappa\exp(\xi t)$, *where* $u(x)\,\varepsilon\,V$. *Then, in the expansion* $u(x)=\Sigma_j c_j x^j$, *the* $c_0, c_1, \cdots, c_\kappa$ *are uniquely determined by* $w\exp(\xi t)$ *and* c_j *for* $j>\kappa$ *is arbitrary and hence may be made zero; furthermore, always* $u(\xi)\neq 0$. *Conversely, an expression* $u(D)t^\kappa\exp(\xi t)$ *always represents an exponential vector of degree* $\leq\kappa$. *The degree is* κ *if and only if* $u(\xi)\neq 0$.

Proof. The symbol $u(\xi)$ denotes of course the constant vector which arises from $u(x)$ by the substitution $x=\xi$. In the same way let $u^{(j)}(\xi)$ denote the vector which arises from the j-th derivative of $u(x)$ with respect to x by the substitution $x=\xi$. The Taylor expansion then gives $u(D)t^\kappa\exp(\xi t)$ $=\exp(\xi t)u(D+\xi)t^\kappa=\exp(\xi t)\Sigma_j(1/j!)u^{(j)}(\xi)D^j t^\kappa=\exp(\xi t)\Sigma_j\binom{\kappa}{j}u^{(j)}(\xi)t^{\kappa-j}$, where $\binom{\kappa}{j}$ is the binomial coefficient $\kappa!/j!(\kappa-j)!$. The lemma follows immediately by comparing this Taylor expansion with the expansion of $w\exp(\xi t)=\exp(\xi t)\Sigma_{j=0}^{\kappa}d_j t^j$ where the d_j are constant vectors and $d_\kappa\neq 0$.

We now return to the system S. The null space N of $A(x)$ is defined as the module of V which consists of the vectors $n(x)$ such that $A(x)n(x)=0$. If $n\,\varepsilon\,N$, clearly $A(D)n(D)t^\kappa\exp(\xi t)=0$ for any κ and ξ. This leads to the following definition.

DEFINITION 1.41. *An exponential solution* $w\exp(\xi t)$ *of* S *is called trivial if there exists an* $n\,\varepsilon\,N$ *such that* $w\exp(\xi t)=n(D)t^\kappa\exp(\xi t)$. *Otherwise, the solution is called non-trivial.*

Observe that, if $w\exp(\xi t)$ is a trivial solution of S of degree κ, $n\,\varepsilon\,N$ can always be chosen in such a way that $w\exp(\xi t)=n(D)t^\kappa\exp(\xi t)$, for there exists some $n\,\varepsilon\,N$ such that $w\exp(\xi t)=n(D)t^\sigma\exp(\xi t)$ for $\sigma\geq\kappa$. If $\sigma>\kappa$, then $n(\xi)=0$ according to Lemma 1.41. Consequently, then $n(x)=(x-\xi)n^*(x)$ where $n^*(x)$ is again a vector of N since *null spaces are clearly always closed modules.* Consequently, $n(D)t^\sigma\exp(\xi t)$ $=n^*(D)(D-\xi)t^\sigma\exp(\xi t)=\sigma n^*(D)t^{\sigma-1}\exp(\xi t)$ where $\sigma n^*(x)\,\varepsilon\,N$. If $\sigma-1>\kappa$ we can repeat this process until we find a vector $n'\,\varepsilon\,N$ such that $w\exp(\xi t)=n'(D)t^\kappa\exp(\xi t)$.

The following theorem establishes the relationship between the exponential solutions of S and the theory of the previous sections. The column space M of A, its invariants Cl (M), Δ, ϵ, r, the non-zero associated primes $\pi_j=x-\xi_j$ $(j=1,\cdots,h)$, and the multiplicities ρ_j of $\epsilon=\Pi_{j=1}^{h}(x-\xi_j)^{\rho_j}$ and l_j of

$\Delta = \Pi_{j=1}^{h}(x-\xi_j)^{l_j}$ are all fixed. Since $\partial_j = 1$ (see Theorem **1.14**), l_j is the P-rank of $M : (x-\xi_j)^{\rho_j}/M$ and $\partial = \Sigma_{j=1}^{h} l_j$ is the P-rank of $\mathrm{Cl}\,(M)/M$.

THEOREM **1.41**. *The exponential vector* $w \exp(\xi t) = u(D)t^{\kappa-1}\exp(\xi t)$ *is an exponential solution of* S, *where* $\kappa \geq 1$, *if and only if* $A(x)u(x) = (x-\xi)^{\kappa}v(x)$. *The solution is non-trivial if and only if* $v(x) \notin M$. *This implies that, if* $w \exp(\xi t)$ *is a non-trivial solution of* S, $x-\xi$ *is a non-zero associated prime of* M, *i.e.,* ξ *is then a root of* ϵ.

Proof. $A(D)u(D)t^{\kappa-1}\exp(\xi t) = \exp(\xi t)\Sigma_{j=0}^{\kappa-1}\binom{\kappa-1}{j}(Au)^{(j)}(\xi)t^{\kappa-1-j}$, where $(Au)^{(j)}(\xi)$ denotes the result of taking the j-th derivative of the vector $A(x)u(x)$ and then substituting $x=\xi$. Hence, $u(D)t^{\kappa-1}\exp(\xi t)$ is a solution of S if and only if $(Au)^{(j)}(\xi) = 0$ for $j = 0, \cdots, \kappa-1$, i.e., if and only if $A(x)u(x) = (x-\xi)^{\kappa}v(x)$. If $v(x) \in M$ the solution would be trivial, for then $v(x) = Af(x)$ and hence $u = (x-\xi)^{\kappa}f + n$, where $n \in N$. This would imply $u(D)t^{\kappa-1}\exp(\xi t) = f(D)(D-\xi)^{\kappa}t^{\kappa-1}\exp(\xi t) + n(D)t^{\kappa-1}\exp(\xi t) = n(D)t^{\kappa-1}\exp(\xi t)$ which proves the triviality. Conversely, if the solution is trivial, there exists an $n \in N$ such that $u(D)t^{\kappa-1}\exp(\xi t) = n(D)t^{\kappa-1}\exp(\xi t)$. From Lemma **1.41** we conclude that then $u(x)$ and $n(x)$ are the same except for terms of degree $\geq \kappa$ and hence $u(x) = n(x) + (x-\xi)^{\kappa}g(x)$. This implies that $A(x)u(x) = (x-\xi)^{\kappa}A(x)g(x) = (x-\xi)^{\kappa}v(x)$ and consequently that $A(x)g(x) = v(x)$, i.e., that $v \in M$. Finally, if $w\exp(\xi t)$ is a non-trivial solution of S, $v(x) \in M : (x-\xi)^{\kappa}$ and $v \notin M$ from which it follows that $M : (x-\xi)^{\kappa} \supset M$ and hence that $M : (x-\xi) \supset M$. According to Definition **1.11**, this means that $x-\xi$ is then a non-zero associated prime of M.

Hence, for the discussion of the non-trivial solutions of S we have to consider only the roots $\xi_1, \cdots, \xi_h$ of ϵ. Consider the sequence $M \subset M : (x-\xi_j) \subset \cdots \subset M : (x-\xi_j)^{\rho_j}$ of **1.1** for some *arbitrarily chosen but fixed root* ξ_j of ϵ and let l_κ be the P-rank of the factor module $M : (x-\xi_j)^{\kappa}/M : (x-\xi_j)^{\kappa-1}$ for $1 \leq \kappa \leq \rho_j$. The P-rank l_κ depends of course on the choice of ξ_j but the index j will always be omitted except for the invariants which were considered in the previous section. Then, $\Sigma_{\kappa=1}^{\rho_j} l_\kappa = l_j$, and, for each $1 \leq \kappa \leq \rho_j$, we can choose l_κ *fixed* vectors $v_{\kappa 1}, \cdots, v_{\kappa l_\kappa}$ of $M : (x-\xi_j)^{\kappa}$ which are P-linearly independent mod $M : (x-\xi_j)^{\kappa-1}$. Naturally, P-linearly independent or dependent stands for linearly independent or dependent with respect to P; consequently, if $\Sigma_{i=1}^{l_\kappa} c_i v_{\kappa i} \in M : (x-\xi_j)^{\kappa-1}$, all the l_κ constants c_i are zero. For each of the l_j vectors $v_{\kappa i}$ we choose some *fixed* vector $u_{\kappa i}$ such that $Au_{\kappa i} = (x-\xi)^{\kappa}v_{\kappa i}$, which can be done since $v_{\kappa i} \in M : (x-\xi)^{\kappa}$. Finally, with each $u_{\kappa i}$ we associate the vector $w_{\kappa i} \in W$ which is uniquely determined by

$w_{\kappa i} \exp(\xi_j t) = u_{\kappa i}(D) t^{\kappa-1} \exp(\xi_j t)$. The following theorem reveals the structure of the non-trivial exponential solutions $w \exp(\xi_j t)$ of S.

THEOREM 1.42. *Let the vectors $v_{\kappa i}$, $u_{\kappa i}$, $w_{\kappa i}$ for $1 \leq \kappa \leq \rho_j$ and $1 \leq i \leq l_\kappa$ be chosen as above. Then, the l_j exponential vectors $w_{\kappa i} \exp(\xi_j t) = u_{\kappa i}(D) t^{\kappa-1} \exp(\xi_j t)$ are l_j, P-linearly independent, non-trivial, exponential solutions of S and the degree of $w_{\kappa i} \exp(\xi_j t)$ is $\kappa - 1$. An arbitrary exponential solution $w \exp(\xi_j t)$ of S is P-linearly dependent on the l_j solutions above and on a trivial solution. If the degree of the solution $w \exp(\xi_j t)$ is $\sigma - 1 \leq \rho_j - 1$, then $w \exp(\xi_j t)$ is P-linearly dependent on the $\Sigma^{\sigma}_{\kappa=1} l_\kappa$ solutions $w_{\kappa i} \exp(\xi_j t)$ for $1 \leq \kappa \leq \sigma$ and $1 \leq i \leq l_\kappa$, and on a trivial solution.*

Proof. Since $A u_{\kappa i} = (x - \xi_j)^{\kappa} v_{\kappa i}$ and since the P-linear independence of $v_{\kappa 1}, \cdots, v_{\kappa l_\kappa} \bmod M : (x - \xi_j)^{\kappa-1}$ implies $v \in M$, the vectors $w_{\kappa j} \exp(\xi_j t)$ are non-trivial exponential solutions of S according to Theorem 1.41. Let $w \exp(\xi_j t) = u(D) t^{\sigma-1} \exp(\xi_j t)$ be an arbitrary exponential solution of degree $\sigma - 1$ of S. Assuming that $\sigma - 1 \geq \rho_j$, i. e., that $\sigma = \rho_j + \gamma$ where $\gamma \geq 1$, we assert that $w \exp(\xi_j t)$ is the sum of an exponential solution of degree $\leq \rho_j - 1$ and a trivial solution of S. We know that $Au = (x - \xi_j)^{\sigma} v$ for some $v \in M : (x - \xi_j)^{\sigma}$ (from Theorem 1.41), and that $M : (x - \xi_j)^{\sigma} = M : (x - \xi_j)^{\rho_j}$ (from Theorem 1.12). Hence, there exists a vector u' such that $Au' = (x - \xi_j)^{\rho_j} v$ and consequently $u = (x - \xi_j)^{\gamma} u' + n$ where $n \in N$. This implies that $u(D) t^{\sigma-1} \exp(\xi_j t) = u'(D)(D - \xi)^{\gamma} t^{\sigma-1} \exp(\xi_j t) + n(D) t^{\sigma-1} \exp(\xi_j t) = c u'(D) t^{\rho_j - 1} \exp(\xi_j t) + n(D) t^{\sigma-1} \exp(\xi_j t)$ where $c \neq 0$ from which the assertion immediately follows. Now, let $\sigma - 1 \leq \rho_j - 1$ and let V_κ be the $m \times l_\kappa$ matrix whose columns are the vectors $v_{\kappa 1}, \cdots, v_{\kappa l_\kappa}$ and let the matrices U_κ and W_κ be similarly defined respectively by the vectors $u_{\kappa 1}, \cdots, u_{\kappa l_\kappa}$ and $w_{\kappa 1}, \cdots, w_{\kappa l_\kappa}$. To say that $w \exp(\xi_j t) = u(D) t^{\sigma-1} \exp(\xi_j t)$ is P-linearly dependent on the solutions $w_{\kappa i} \exp(\xi_j t)$ for $1 \leq \kappa \leq \sigma$ and $1 \leq i \leq l_\kappa$ means of course that $w \exp(\xi_j t) = (\Sigma^{\sigma}_{\kappa=1} W_\kappa c_\kappa) \exp(\xi_j t)$ where the c_κ's are constant vectors with l_κ components. A simple computation shows that this is equivalent to saying that $u(D) t^{\sigma-1} \exp(\xi_j t) = \Sigma^{\sigma}_{\kappa=1} (U_\kappa c_\kappa)(D)(t^{\kappa-1} \exp(\xi_j t))$, where $(U_\kappa c_\kappa)(D)$ is the operator vector which arises from the vector $U_\kappa c_\kappa$ by the substitution $x = D$. In order to prove the existence of the σ vectors c_κ we observe that, according to Theorem 1.41, $Au = (x - \xi_j)^{\sigma} v$ where $v \in M : (x - \xi_j)^{\sigma}$. Since $v_{\sigma 1}, \cdots, v_{\sigma l_\sigma}$ are σ vectors of $M : (x - \xi_j)^{\sigma}$ which are P-linearly independent mod $M : (x - \xi_j)^{\sigma-1}$ and since σ is the P-rank of the factor module $M : (x - \xi_j)^{\sigma} / M : (x - \xi_j)^{\sigma-1}$, there exists a constant vector c with l_σ components such that $v = V_\sigma c + v'$

where $v' \varepsilon M : (x-\xi_j)^{\sigma-1}$. Hence, $(x-\xi_j)^{\sigma}v' = (x-\xi_j)Av''$ which implies that $A(u-(x-\xi_j)v'') = (x-\xi_j)^{\sigma}V_{\sigma}c$. Since also $AU_{\sigma}c = (x-\xi_j)^{\sigma}V_{\sigma}c$, we conclude that $u = (x-\xi_j)v'' + U_{\sigma}c + n$, where $n \varepsilon N$. Consequently, $u(D)t^{\sigma-1}\exp(\xi_j t) = v''(D)\exp(\xi_j t)(Dt^{\sigma-1}) + (U_{\sigma}c)(D)t^{\sigma-1}\exp(\xi_j t) + n(D)t^{\sigma-1}\exp(\xi_j t)$. The last term is a trivial solution of S and the middle term is a solution which is P-linearly dependent on the solutions $u_{\sigma i}(D)t^{\sigma-1}\exp(\xi_j t)$ for $1 \leq i \leq l_{\sigma}$ of S. Hence, the first term is a solution of S of degree at most $\sigma-2$ and is not present if $\sigma-1=0$. Hence, if $\sigma-1=0$, the solution $w\exp(\xi_j t)$ is P-linearly dependent on the solutions stated in Theorem 1.42. If $\sigma-1>0$, we can first make the induction hypothesis that the P-linear dependence of $w\exp(\xi_j t)$ on the solutions stated in Theorem 1.42 is proved for degrees $0, 1, \cdots, \sigma-2$. For degree $\sigma-1$ the same statement then follows immediately from the above remark about the degree of the first term. There remains to be proved that the degree of $w_{\kappa i}\exp(\xi_j t)$ is $\kappa-1$ and that the l_j solutions of Theorem 1.42 are P-linearly independent. Let $u_{\kappa i}(\xi_j)$ be the constant vector which arises from $u_{\kappa i}$ by the substitution $x=\xi_j$. *We assert that the l_{κ} vectors $u_{\kappa 1}(\xi_j), \cdots, u_{\kappa l_{\kappa}}(\xi_j)$ are P-linearly independent for each $1 \leq \kappa \leq \rho_j$.* To prove this assertion, suppose that $U_{\kappa}(\xi_j)c=0$ where c is a constant vector with l_{κ} components and where the matrix $U_{\kappa}(\xi_j)$ arises from $U_{\kappa}(x)$ by the substitution $x=\xi_j$. Then $U_{\kappa}(x)c = (x-\xi_j)u'$ and hence $(x-\xi_j)Au' = (x-\xi_j)^{\kappa}V_{\kappa}c$, which implies that $Au' = (x-\xi_j)^{\kappa-1}V_{\kappa}c$. This means that the vector $V_{\kappa}c \varepsilon M : (x-\xi_j)^{\kappa-1}$ and hence that $c=0$ which proves the assertion. The assertion implies that $u_{\kappa i}(\xi_j) \neq 0$ and hence that the degree of $u_{\kappa i}(D)t^{\kappa-1}\exp(\xi_j t)$ is $\kappa-1$ according to Lemma 1.41. Furthermore,

$$u_{\kappa i}(D)t^{\kappa-1}\exp(\xi_j t) = \exp(\xi_j t)\sum_{g=0}^{\kappa-1}\binom{\kappa-1}{g}u_{\kappa i}^{(g)}(\xi_j)t^{\kappa-1-g}$$

(see the proof of Lemma 1.41) and hence the coefficient of $t^{\kappa-1}$ is the vector $u_{\kappa i}(\xi_j)$. From the above assertion we then conclude that, for each κ, the l_{κ} vectors $u_{\kappa}(D)t^{\kappa-1}\exp(\xi_j t)$ are P-linearly independent and, from the fact that the degree of $u_{\kappa}(D)t^{\kappa-1}\exp(\xi_j t)$ is $\kappa-1$, we conclude that all the l_j vectors of Theorem 1.42 are P-linearly independent.

With regard to the trivial solutions of S, let $n_1, \cdots, n_{m-r}$ be $m-r$ linearly independent generators of N. Clearly, a trivial solution is then P-linearly dependent on the infinite number of P-linearly independent trivial solutions $n_i(D)t^{\kappa}\exp(\xi_j t)$ for $1 \leq i \leq m-r$ and arbitrary κ.

Remark 1.41. If A is non-singular, $r=m$ and S has no trivial solutions. Hence, Theorem 1.42 contains the usual statement about the system S.

namely that, if A is a square non-singular matrix, the number of P-linearly independent exponential solutions $w \exp(\xi_j t)$ of S is equal to the multiplicity l_j of ξ_j as a root of Δ (see [1], Chapter 5).

1.5. Computation of the solution of S.

Remark 1.51. The purpose of this section is to show how the solutions of S can be computed by means of the methods of the previous section. The difference between this method and the classical way of computing the solutions (see [1] pp. 168 and 179) is that the matrix A does not have to be restricted to a non-singular square matrix.

The notation will be the same as in **1.4.** Hence A is again an $m \times s$ matrix of rank r whose elements are polynomials of $P[x]$, and the invariants Δ, ϵ, l_j, ρ_j, ξ_j, Cl (M) etc. all refer to the column space M of A. Again, the letter c denotes elements of P and N is the null space of A.

DEFINITION 1.51. *An $s \times r$ matrix B will be called a ξ_j-multiplier of multiplicity σ of A if $AB = \phi C$, where ϕ is a polynomial of $P[x]$ which has ξ_j as a root of multiplicity σ and where the r columns of C generate* Cl (M).

Since $M \subseteq$ Cl (M) an $r \times s$ matrix K exists such that $A = CK$ (designated equation m_1) which implies that $ABK = \phi A$ and hence that $BK + F = \phi I$ (designated equation m_2), where F is an $s \times s$ matrix whose columns are elements of N and where I is the $s \times s$ unit matrix. The following theorem properly states how solutions of S can be computed if A is an arbitrary $m \times s$ matrix of rank r.

THEOREM 1.51. *Let B be an ξ_j-multiplier of multiplicity $\sigma \geq \rho_j$ of A. Let the $m \times r$ matrix $U_\kappa(t)$ for $0 \leq \kappa \leq \sigma - 1$ be defined by $U_\kappa \exp(\xi_j t) = B(D) t^\kappa \exp(\xi_j t)$ where $B(D)$ arises from B by the substitution $x = D$. Then, the columns of $U_\kappa \exp(\xi_j t)$ are exponential solutions of S. Conversely, every exponential solution $w \exp(\xi_j t)$ of S is P-linearly dependent on these columns and on a trivial solution.*

Proof.

$$A(D) U_\kappa \exp(\xi_j t) = A(D) B(D) t^\kappa \exp(\xi_j t) = C(D) \phi(D) t^\kappa \exp(\xi_j t).$$

Since $\phi(D) = \psi(D)(D - \xi_j)^\sigma$ and $\sigma > \kappa$, $C(D)\phi(D) t^\kappa \exp(\xi_j t) = 0$ and hence every column of $U_\kappa \exp(\xi_j t)$ is an exponential solution of S for $0 \leq \kappa \leq \sigma - 1$. To prove the converse, observe that, according to Theorem 1.42, every exponential solution of S is the sum of an exponential solution

of degree $\leq \rho_j - 1$ and a trivial solution. Hence we only have to assume that $w \exp(\xi_j t) = u(D) t^{\rho_j - 1 - \gamma} \exp(\xi_j t)$ is a solution of degree $\rho_j - 1 - \gamma$ of S for $0 \leq \gamma \leq \rho_j - 1$. Then we know that $Au = (x - \xi_j)^{\rho_j - \gamma} v$ where $v \in \mathrm{Cl}(M)$ and hence that $v = Cv'$. Equation m_1 then gives $CKu = (x - \xi_j)^{\rho_j - \gamma} Cv'$ and, since the r columns of C, as generators of $\mathrm{Cl}(M)$, must be linearly independent with respect to $P[x]$, we conclude that $Ku = (x - \xi_j)^{\rho_j - \gamma} v'$. Consequently $BKu = B(x - \xi_j)^{\rho_j - \gamma} v'$ and hence, from equation m_2, $\phi u = B(x - \xi_j)^{\rho_j - \gamma} v' + Fu$. Since $\phi = (x - \xi_j)^{\sigma} \psi$ and $\sigma \geq \rho_j$, $Fu = (x - \xi_j)^{\rho_j - \gamma} n$, where $n \in N$, since $Fu \in N$ and a null space is always a closed module. We conclude that, if $\sigma = \rho_j + h$ for $h \geq 0$, $(x - \xi_j)^{h+\gamma} \psi u = Bv' + n$. This gives rise to the relationship $u(D)\psi(D)(D - \xi_j)^{h+\gamma} t^{\rho_j + h - 1} \exp(\xi_j t) = B(D)v'(D) t^{\rho_j + h - 1} \exp(\xi_j t) + n(D) t^{\rho_j + h - 1} \exp(\xi_j t)$. The last term of this relationship is a trivial solution of S. The first term is P-linearly dependent on the columns of $U_\kappa \exp(\xi_j t)$ for $0 \leq \kappa \leq \rho_j + h - 1 = \sigma - 1$ as is seen by first expanding $v'(D) = \Sigma_i c_i D^i$ and then applying this operator to $t^{\rho_j + h - 1} \exp(\xi_j t)$. For the left-hand side we write $u(D)\psi(D)(D - \xi_j)^{h+\gamma} t^{\rho_j + h - 1} \exp(\xi_j t) = cu(D)\psi(D) t^{\rho_j - \gamma - 1} \exp(\xi_j t)$, where $c \neq 0$. Since $\psi(D) = \Sigma d_i (D - \xi_j)^i$, where $d_i \in P$ and $d_0 \neq 0$ because $\psi(\xi_j) \neq 0$, application of the operator $\psi(D)$ to $t^{\rho_j - \gamma - 1} \exp(\xi_j t)$ shows that $cu(D)\psi(D) t^{\rho_j - \gamma - 1} \exp(\xi_j t) = dw \exp(\xi_j t) + w' \exp(\xi_j t)$ where d is the non-zero element $d_0 c$ of P and where $w \exp(\xi_j t)$ is the given exponential solution while $w' \exp(\xi_j t)$ is an exponential solution of degree $< \rho_j - \gamma - 1$ of S. Hence, if the degree of $w \exp(\xi_j t)$ is zero, i. e., if $\rho_j - 1 - \gamma = 0$, the term $w' \exp(\xi_j t)$ does not occur and the theorem is proved. For arbitrary degree $\rho_j - 1 - \gamma$, we first make the induction hypothesis that the theorem is proved for degrees $0, 1, \cdots, \rho_j - 2 - \gamma$, and the theorem then follows for degree $\rho_j - 1 - \gamma$ from the remark about the degree of $w' \exp(\xi_j t)$.

There always exist matrices B which are ξ_j-multipliers of A of multiplicity ρ_j simultaneously for all roots ξ_j of ϵ, for, if C is any matrix whose columns generate $\mathrm{Cl}(M)$, then ϵC is a matrix whose columns are vectors of M. Hence, there exists a matrix B such that $AB = \epsilon C$ and this B clearly satisfies the requirement.

Remark **1.** 52. In the classical method of computing the solutions of S it is assumed that A is a square non-singular matrix and hence its adjoint H is then well defined. The substitution $x = \xi_j$ is made in the derivatives of order $0, 1, \cdots, l_j - 1$ of $\exp(xt) H(x)$ with respect to x (see [1], pp. 168 and 179). However, H is then a ξ_j-multiplier of A of multiplicity $l_j \geq \rho_j$ simultaneously for all roots ξ_j of ϵ, because $AH = \Delta I$ and the columns of

the unit matrix I generate the whole vector space V, which is Cl (M), since M is dense in this case. Furthermore

$$B(D)t^\kappa \exp(\xi_j t) = \exp(\xi_j t)\Sigma_{i=0}^{\kappa}\binom{\kappa}{\kappa-i}B^{(i)}(\xi_j)t^{\kappa-i}$$
$$= \exp(\xi_j t)[(D_x + t)^\kappa B(x)]_{x=\xi_j} = [D_x^\kappa \exp(xt)\cdot B(x)]_{x=\xi_j}$$

which shows that the classical statement, mentioned above is contained in Theorem **1.** 51.

1.6. Trivial solutions in terms of the row space. According to Theorem **1.** 42 the properties of the solution of S are determined by two column space invariants, namely Δ and ϵ, and by the trivial solutions. The trivial solutions are defined in terms of the null space N which is not an invariant of the column space M. The rows of A, however, generate the row space M' which is a module of the m-dimensional vector space V' consisting of the row vectors whose m components are elements of $P[x]$. Since the Δ and ϵ of M and M' can be computed in the same way from the sub-determinants of A, the Δ and ϵ of M and M' are the same. Hence, we can as well say that the properties of S are determined by two row space invariants and by the trivial solutions. We shall now show how the trivial solutions could have been defined invariantly in terms of the row space.

Remark **1.** 61. The purpose of this section is to give the basis for the notion of trivial solution for a system of partial differential equations which will be discussed in MSV. Since, if the scalar domain is $\mathfrak{s} = P[x_1, \cdots, x_n]$, the invariants Δ and ϵ can not be computed any more as sub-determinants of A, a column space and a row space of the same matrix may then very well have different invariants. It will be shown in MSV that if $\mathfrak{s} = P[x_1, \cdots, x_n]$, the invariants of the row space M', and not of the column space M, are essential for the system of partial differential equations determined by the matrix. In particular, trivial solutions then have to be defined in terms of M'. It will become clear from MSV how the theory of **1.4** could have been developed, less simply but yet completely, in terms of the row space M'. This is no wonder since, if $\mathfrak{s} = P[x]$, the fact that the invariants of M and M' can be computed by means of sub-determinants of A, implies that M and M' then have the same invariants.

The system S is given by m differential operators which arise from the m rows of A by the substitution $x = D$. Since every vector of M' is a linear combination of the rows of A, the system of differential equations $M'(D)$, which is given by the infinite number of differential operators which arise

from all the rows of M' by the substitution $x = D$, has the same solutions as S. If Cl $(M')(D)$ arises from Cl (M') in the same way as $M'(D)$ from M', then, since $M' \subseteq$ Cl (M'), the solutions of Cl $(M')(D)$ are contained among those of $M'(D)$, i. e., among those of S. The following theorem shows how the trivial solutions of S could have been defined intrinsically in terms of M'.

THEOREM **1. 61.** *An exponential solution* $w \exp(\xi t)$ *of* S *is a trivial solution of* S *in the sense of Definition* **1. 41**, *if and only if it is a solution of* Cl $(M')(D)$.

Proof. Let $w \exp(\xi t)$ be a trivial exponential solution of S, i. e., let $w \exp(\xi t) = n(D) t^\kappa \exp(\xi t)$ where $An = 0$. Let $v'(D) \in$ Cl $(M')(D)$, i. e., v' is a row vector of Cl (M'). Since ϵ is the elementary divisor of M', we conclude that $\epsilon v' \in M'$ and hence there exists a row vector $s' \in V'$ such that $s'A = \epsilon v'$. This implies that $\epsilon v'n = 0$ and hence that $v'n = 0$ from which it follows that $v'(D) w \exp(\xi t) = v'(D) n(D) t^\kappa \exp(\xi t) = 0$. Consequently, every trivial solution is a solution of Cl $(M')(D)$. Conversely, if $w \exp(\xi t) = u(D) t^\kappa \exp(\xi t)$ is a solution of Cl $(M')(D)$, it is also a solution of $M'(D)$, and hence all that remains to be proved is that this solution is trivial. Let F be any matrix whose rows generate Cl (M'); then, since $F(D) u(D) t^\kappa \exp(\xi t) = 0$, we have $Fu = (x - \xi)^\kappa v$. Since the rows of F generate a closed module, i. e., a module whose Δ is a non-zero element of P, the determinantal factor of highest rank of F, i. e., Δ, is an element of P. Hence, the columns of F also generate a closed module, which implies the existence of a vector $s \in V$ such that $Fs = v$. Consequently, according to Theorem **1. 41**, $u(D) t^\kappa \exp(\xi t)$ is a trivial solution of $F(D) y = 0$, i. e., $u(D) t^\kappa \exp(\xi t) = n(D) t^\kappa \exp(\xi t)$ where $Fn = 0$. Since the rows of A are vectors of M' and hence of Cl (M'), we can take for F a matrix whose first m rows are those of A. We then conclude that $An = 0$ and hence that $w \exp(\xi t)$ is a trivial solution of S.

Part II. Structure of Modules.

2. 1. Operations and the notion of primary. Let $\mathfrak{s}$ be a commutative ring with unit element and zero element ω. The elements of $\mathfrak{s}$, called scalars, will be denoted by lower case Greek letters and the ideals of $\mathfrak{s}$ by lower case German letters. Let V be an additive abelian group with zero element 0 for which $\mathfrak{s}$ is the left operator domain and for which the unit element of $\mathfrak{s}$ is the unit operator. The elements of V, called vectors, will be denoted by lower case Roman letters and the $\mathfrak{s}$-modules of V, called modules, by capital Roman letters. Consequently, a module is a subset of V which is closed under

scalar multiplication and vector-subtraction. We assume furthermore that *for the modules of V the ascending chain condition holds*, i. e., that each module has a finite number of generators. For the terms used so far, see [9] and [11]. In accordance with [12], p. 690, we call V *a Noetherian* vector space with $\mathfrak{s}$ as scalar domain, or more briefly, *a Noetherian vector space over* $\mathfrak{s}$. If, without further explanation, a lower case Greek letter, a lower case German letter, a lower case Roman letter or a capital Roman letter is used, these letters will always respectively denote a scalar, an ideal, a vector and a module.

Remark **2.** 11. The assumptions made in the previous paragraph are indicated by [0, I] in [5]. The vector space of Part I is an example of a Noetherian vector space over $P[x]$ as scalar domain.

For the consistency of the following definitions, see [5] Section 2 or [6] Section 2.

The intersection $M \cap L$ is the module which is the set theoretic intersection of M and L. It is convenient for **2.3** to say, as in [5], that a void collection of modules has V as intersection. The sum (M, L) is the module which is generated by the set theoretic union $M \cup L$ and hence consists of the vectors $v = m + l$ where $m \in M$ and $l \in L$. In the same way, $(v_1, \cdots, v_s)$ denotes the module generated by $v_1, \cdots, v_s$ and $(\alpha_1, \cdots, \alpha_s)$ the ideal generated by $\alpha_1, \cdots, \alpha_s$. The *quotient* $M : L$ is the ideal of $\mathfrak{s}$ which consists of the scalars α such that $\alpha l \in M$ for all $l \in L$. Here, L can be an arbitrary set of vectors, not necessarily a module. The *quotient* $M : \mathfrak{b}$ is the module which consists of all vectors v such that $\beta v \in M$ for all $\beta \in \mathfrak{b}$. Again, $\mathfrak{b}$ can be any set of scalars, not necessarily an ideal. The *product* $\mathfrak{b}M$ is the module which is generated by all products βm where $\beta \in \mathfrak{b}$ and $m \in M$. Again, $\mathfrak{b}$ may be an arbitrary set of scalars. The rules which combine intersections, sums, quotients and products are the same as for ideals (see [5] Section 2 or [6] Section 2).

DEFINITION **2.** 11. *The fundamental ideal* $\mathfrak{f}$ *of* M *is the quotient* $M : V$. *The radical* $\mathfrak{r}$ *of* M *is the radical of* $\mathfrak{f}$.

The radical of an ideal is as usual the ideal consisting of the scalars of which a power lies in the ideal.

Remark **2.** 12. The term "fundamental ideal" in definition **2.** 11 means the same as the terms "shadow" in [5] and "essential ideal" in [6]. The term "shadow" is dropped since it does not indicate that the structure of $\mathfrak{f}$ is of "fundamental" importance for M. The reason for changing the terminology of [6] is to keep the German letter $\mathfrak{e}$ free to denote "elementary divisor."

From now on, the annihiliating ideal of M will always be denoted by $\mathfrak{a}$, i. e., $\mathfrak{a}(M) = 0 : M$. Since $M_1 \subseteq M_2$ implies $\mathfrak{f}(M_1) \subseteq \mathfrak{f}(M_2)$ and $\mathfrak{r}(M_1) \subseteq \mathfrak{r}(M_2)$, it is always true that $\mathfrak{a}(V) \subseteq \mathfrak{f}(M)$ and $\mathfrak{r}(0) \subseteq \mathfrak{r}(M)$. The following definition and theorem occur in [5], Section 4, and in [6], Section 4.

DEFINITION **2.12.** *A module Q is called primary if $\alpha v \in Q$ and $\alpha \notin \mathfrak{r}(Q)$ implies $v \in Q$.*

THEOREM **2.11.** *The fundamental ideal of a primary module is a primary ideal and hence its radical is a prime ideal. A module Q is $\mathfrak{p}$-primary, i. e., is primary and has $\mathfrak{p}$ as radical, if and only if the following three conditions are satisfied*: 1) *$\alpha v \in Q$ and $v \notin Q$ implies $\alpha \in \mathfrak{p}$*; 2) *$\mathfrak{f}(Q) \subseteq \mathfrak{p}$*; 3) *$\alpha \in \mathfrak{p}$ implies that, for some integer κ, $\alpha^\kappa \in \mathfrak{f}(Q)$.*

Since, if Q is $\mathfrak{p}$-primary, $\mathfrak{p} \supseteq \mathfrak{f}(Q) \supseteq \mathfrak{a}(V)$ and since the ascending chain condition holds in $\mathfrak{F}/\mathfrak{a}(V)$ (see [5], Theorem 4), there is a smallest integer ρ such that $\mathfrak{p}^\rho \subseteq \mathfrak{f}(Q)$. This ρ is called the *exponent of* Q. If Q is $\mathfrak{p}$-primary and $\mathfrak{f}(Q)$ is prime, i. e., if $\mathfrak{f}(Q) = \mathfrak{p}$, and $\rho = 1$, then Q is called *a prime module* or *$\mathfrak{p}$-prime*. The following lemma is proved in [5], Section 4.

LEMMA **2.11.** *If Q is $\mathfrak{p}$-primary, $\mathfrak{b} \not\subseteq \mathfrak{f}(Q)$ and $M \not\subseteq Q$, then $Q : \mathfrak{b}$ is a $\mathfrak{p}$-primary module and $Q : M$ a $\mathfrak{p}$-primary ideal.*

If $M \subseteq L$ and if M is $\mathfrak{p}$-primary when M is considered as a sub-module of L instead of V, then M is said to be *$\mathfrak{p}$-primary in L*. Since the fundamental ideal of M as a sub-module of L is $M : L$ and not $M : V$, it may be that M is $\mathfrak{p}$-primary in L but not in V. However, if M is $\mathfrak{p}$-primary in V, it is certainly $\mathfrak{p}$-primary in L (if $M \neq L$) as is implied by Lemma **2.12**, which follows from the three conditions of Theorem **2.11**.

LEMMA **2.12.** *Let Q be $\mathfrak{p}$-primary and let L be any module, not contained in Q. Then $Q \cap L$ is $\mathfrak{p}$-primary in L and the $\mathfrak{p}$-primary fundamental ideal of $Q \cap L$ as a sub-module of L is $Q : L$.*

Remark **2.13.** Lemma **2.12**, which does not occur in [5] and [6], will be used frequently in that part of the theory of polynomial matrices in which the imbedding vector space V has to be replaced by an isolated component of the row space or column space of the matrix. See also Lemma **2.51** of this paper.

2.2. Noether decompositions.

DEFINITION **2.21.** *An intersection $M = Q_1 \cap \cdots \cap Q_h$ is called a Noether decomposition (N. D. for short) of M if the following three conditions are*

satisfied: 1) Q_j *is* $\mathfrak{p}_j$*-primary for* $j = 1, \cdots, h$; 2) *the prime ideals* $\mathfrak{p}_1, \cdots, \mathfrak{p}_h$ *are all different*; 3) *the intersection is irredundant.*

Remark **2.** 21. The term "Noether decomposition" is used in preference to the term "Lasker reduction" of [5] to conform to the expression Noetherian vector space.

The following theorem is proved in [5], Section 6, and in [6], Sections 4 and 5.

THEOREM **2.** 21. *Every module* M *has an N. D. The radicals of the primary components are the same for all N. D.'s of* M *and are called the associated primes* (*a. p. for short*) *of* M. *The intersection of the a. p.'s of* M *is the radical of* M.

A primary component of M *is a primary module which occurs in some N. D. of* M *and a complete set of primary components of* M *is a set of primary components which together form an N. D. of* M. The intersection of a subset of a complete set of primary components of M is called a component of M. From Lemma **2.** 12 the following lemma can be easily derived.

LEMMA **2.** 21. *Let* $M = Q_1 \cap \cdots \cap Q_h$ *be an N. D. of* M, *where* Q_j *is* $\mathfrak{p}_j$*-primary for* $1 \leq j \leq h$. *Let* C *be the component* $C = Q_1 \cap \cdots \cap Q_\kappa$, *where* $1 \leq \kappa \leq h$, *and let* $Q^*_{\kappa+i} = C \cap Q_{\kappa+i}$, *where* $1 \leq i \leq h - \kappa$. *Then* $M = Q^*_{\kappa+1} \cap \cdots \cap Q^*_h$ *is an N. D. of* M *as a submodule of* C *and the a. p.'s of* M *as a submodule of* C *are* $\mathfrak{p}_{\kappa+1}, \cdots, \mathfrak{p}_h$.

Remark **2.** 22. Lemma **2.** 21 does not occur in [5] and [6] and is important for the same reason as Lemma **2.** 12.

A component of M is called an *isolated component* if every a. p. of M which is contained in an a. p. of that component is itself an a. p. of that component. An isolated component of M is uniquely determined by its a. p.'s as is proved in [5], Section 6, and in [6], Section 5. A module is called univalent ("monotypic" in [5]) if its radical is a maximal ideal of $\mathfrak{s}$. Since a univalent module can clearly have only one a. p., namely its own radical, a univalent module is always primary. Finally, an arbitrary module and an ideal $\mathfrak{c}$ are said to be *relatively prime* if $M : \mathfrak{c} = M$. It is shown on page 254 of [5] that M and $\mathfrak{c}$ are relatively prime if and only if no a. p. of $\mathfrak{c}$ is contained in an a. p. of M.

Remark **2.** 23. According to the above, if $\mathfrak{p}$ is a prime ideal, $M : \mathfrak{p} \supset M$ (the symbol $\supset$ again denotes *proper* inclusion) if and only if $\mathfrak{p}$ is contained in one of the a. p.'s of M. Since in the scalar domain $P[x]$ of Part I all

prime ideals are maximal, in that particular case $M : \mathfrak{p} \supset M$ if and only if $\mathfrak{p}$ is an a. p. of M. Hence, the non-zero a. p.'s of Part I are those a. p.'s of an N. D. of M which are different from the zero ideal.

2. 3. Closures of a module. Throughout this section, let $\mathfrak{C}$ denote a multiplicatively closed subset of $\mathfrak{S}$ which does not contain ω.

DEFINITION **2. 31.** *The $\mathfrak{C}$-closure $M(\mathfrak{C})$ of a module M is the module which consists of the vectors v for which there exists a scalar $\gamma \in \mathfrak{C}$ such that $\gamma v \in M$. If $M(\mathfrak{C}) = M$, we say that M is $\mathfrak{C}$-closed, and if $M(\mathfrak{C}) = V$ that M is $\mathfrak{C}$-dense.*

Clearly, $M \subseteq M(\mathfrak{C})$ and, if $M_1 \subseteq M_2$, then $M_1(\mathfrak{C}) \subseteq M_2(\mathfrak{C})$.

Remark **2. 31.** If, in the scalar domain $P[x]$ of Part I, we take for $\mathfrak{C}$ the set of all the non-zero scalars, the $\mathfrak{C}$-closure becomes the closure Cl (M) of Definition **1. 12.** This particular closure occurs in [6] Section 3 for general m-dimensional vector spaces. As will be shown in MSV, the n different closures of a polynomial matrix in n variables, as defined in [13] Section 3, are closures in the sense of Definition **2. 31.**

The important properties of closures are enumerated in the following lemma whose proof can be found in [5] Section 7.

LEMMA **2. 31.** *The $\mathfrak{C}$-closure $M(\mathfrak{C})$ is the smallest $\mathfrak{C}$-closed module which contains M and is the isolated component of M whose a. p.'s contain no element of $\mathfrak{C}$. Furthermore, $[M_1 \cap M_2](\mathfrak{C}) = M_1(\mathfrak{C}) \cap M_2(\mathfrak{C})$, from which it follows that the intersection of a finite number of $\mathfrak{C}$-closed modules is $\mathfrak{C}$-closed, that the intersection of a finite number of $\mathfrak{C}$-dense modules is $\mathfrak{C}$-dense, and the irredundant intersection of a $\mathfrak{C}$-closed module and a $\mathfrak{C}$-dense module is neither $\mathfrak{C}$-closed nor $\mathfrak{C}$-dense.*

By considering all possible $\mathfrak{C}$-closures of M we obtain all possible isolated components of M (see [5] Theorem 16) together with V. Namely, if every a. p. of M has an element in common with $\mathfrak{C}$, then M is $\mathfrak{C}$-dense, according to the convention that the intersection of a void collection of modules is V. If $\mathfrak{C}$ is the complement of a prime ideal $\mathfrak{p}$ of $\mathfrak{S}$, we write $M(\mathfrak{p})$, $\mathfrak{p}$-closure and $\mathfrak{p}$-closed instead of $M(\mathfrak{C})$, $\mathfrak{C}$-closure and $\mathfrak{C}$-closed, following the custom of [5] for modules and [14] for ideals. Hence, $M(\mathfrak{p})$ is the isolated component of M whose a. p.'s are contained in $\mathfrak{p}$. It follows that M is $\mathfrak{p}$-closed if and only if all a. p.'s of M are contained in $\mathfrak{p}$ and is $\mathfrak{p}$-dense if and only if none of its a. p.'s are contained in $\mathfrak{p}$. In particular, a $\mathfrak{p}$-primary module is $\mathfrak{p}$-closed.

Remark **2. 32.** If V is the vector space of Part I, Cl $(M) = M((\omega))$,

where (ω) is the zero ideal of $P[x]$; and closed and dense then mean (ω)-closed and (ω)-dense. Hence, in that case, a module M is dense if and only if (ω) is not an a. p. of M. If, however, M is not dense, (ω) is an a. p. of M and Cl (M) is then the isolated primary component of M whose a. p. is (ω). The occurrence of the zero ideal as an a. p. does not happen in ideal theory since ideals of integral domains, considered as modules, are always (ω)-dense.

We now discuss the notion of quotient module which is related to that of $\mathfrak{C}$-closed module and is needed for the theory of length of a primary module. The *quotient module* $V_{\mathfrak{C}}$ consists of classes of ordered pairs $\{v, \gamma\}$ where $v \in V$, $\gamma \in \mathfrak{C}$ and two pairs $\{v_1, \gamma_1\}$ and $\{v_2, \gamma_2\}$ belong to the same class if $\gamma_2 v_1 = \gamma_1 v_2$. According to [5] Section 13, addition and scalar multiplication can be defined in the usual way in $V_{\mathfrak{C}}$, which makes $V_{\mathfrak{C}}$ into a Noetherian vector space over $\mathfrak{s}_{\mathfrak{C}}$ as scalar domain. Here, $\mathfrak{s}_{\mathfrak{C}}$ arises from $\mathfrak{s}$ as $V_{\mathfrak{C}}$ arises from V. By forming quotients, the module $M_{\mathfrak{C}} = T(M)$ of $V_{\mathfrak{C}}$ corresponds to the module M of V. In the same way, the ideal $\mathfrak{b}_{\mathfrak{C}} = A(\mathfrak{b})$ of $\mathfrak{s}_{\mathfrak{C}}$ corresponds to the ideal $\mathfrak{b}$ of $\mathfrak{s}$. The following lemma is proved in [5] Section 13.

LEMMA **2.32.** *The correspondence T is a one-to-one correspondence between the $\mathfrak{C}$-closed modules of V and the modules of $V_{\mathfrak{C}}$ and A is a one-to-one correspondence between the $\mathfrak{C}$-closed ideals of $\mathfrak{s}$ and the ideals of $\mathfrak{s}_{\mathfrak{C}}$. Both these correspondences are isomorphisms with respect to the quotient operation :, the intersection operation $\cap$, and the inclusion relation $\subset$. It follows that, if M is $\mathfrak{C}$-closed, $\mathfrak{f}(T(M)) = A(\mathfrak{f}(M))$ and $\mathfrak{r}(T(M)) = A(\mathfrak{r}(M))$. Furthermore M is $\mathfrak{p}$-primary if and only if $T(M)$ is $A(\mathfrak{p})$-primary, in which case M and $T(M)$ have the same exponents.*

Again, if $\mathfrak{C}$ is the complement of a prime ideal $\mathfrak{p}$ of $\mathfrak{s}$, we write $V_{\mathfrak{p}}$, $\mathfrak{s}_{\mathfrak{p}}$ etc. instead of $V_{\mathfrak{C}}$, $\mathfrak{s}_{\mathfrak{C}}$.

2.4. Length of primary modules. The following definition occurs for ideals in [10] Section 23.

DEFINITION **2.41.** *A sequence $Q_0 \subset Q_1 \subset \cdots \subset Q_{l-1} \subset V$ is called a $\mathfrak{p}$-primary composition sequence from Q_0 to V of length l if* 1) *Q_i is $\mathfrak{p}$-primary for $0 \leq i \leq l-1$ and* 2) *it is impossible to insert a $\mathfrak{p}$-primary module Q between any two consecutive members $Q_i \subset Q_{i+1}$ such that $Q_i \subset Q \subset Q_{i+1}$.*

Let Q be a $\mathfrak{p}$-primary module. Since Q is $\mathfrak{p}$-closed, $T(Q)$ is a primary $A(\mathfrak{p})$ module of $V_{\mathfrak{p}}$. Since $A(\mathfrak{p})$ is a maximal prime ideal of $\mathfrak{s}_{\mathfrak{p}}$ (see [5] Section 14), $T(Q)$ is even univalent and hence the factor module $V_{\mathfrak{p}}/T(Q)$ is a Noetherian vector space over $\mathfrak{s}_{\mathfrak{p}}$ as scalar domain whose zero module is univalent. This implies (see [5] Section 11) that $V_{\mathfrak{p}}/T(Q)$ has an ordinary

composition sequence whose factors (i. e., the factor modules formed by consecutive elements) are all isomorphic with $\mathfrak{S}_\mathfrak{p}/A(\mathfrak{p})$. These isomorphisms are operator isomorphisms with respect to $\mathfrak{S}_\mathfrak{p}$ as operator domain. The theory of composition sequences, applied to $V_\mathfrak{p}/T(Q)$, then gives, by the use of Lemma **2.** 32, the following theorem.

THEOREM **2.** 41. *For every $\mathfrak{p}$-primary module Q we can draw a $\mathfrak{p}$-primary composition sequence $Q \subset Q_1 \subset \cdots \subset Q_{l-1} \subset V$ from Q to V. All these sequences have the same length l, called the length of Q. Any sequence $Q \subset H_1 \subset \cdots \subset H_s \subset V$ where the H_i's are $\mathfrak{p}$-primary can be refined to a $\mathfrak{p}$-primary composition sequence. If Q is univalent, a $\mathfrak{p}$-primary composition sequence from Q to V is an ordinary composition sequence, i. e., no module, $\mathfrak{p}$-primary or not, can be inserted between two consecutive elements. In that case the factors are all operator isomorphic with $\mathfrak{S}/\mathfrak{p}$ with respect to $\mathfrak{S}$ as operator domain.*

Remark **2.** 41. In Theorem **1.** 13 it was shown that, for the vector space of Part I, a composition sequence can be drawn from any module M, *primary or not,* to any other module N where $M \subseteq N \subseteq \mathrm{Cl}\,(M)$. This is caused by the fact that all non-zero prime ideals of $P[x]$ are maximal and hence all primary components of M, except $\mathrm{Cl}\,(M)$, are univalent (see [5] Theorem 24).

2. 5. Elementary divisors and lengths of arbitrary modules. Let $\mathfrak{p}$ be an a. p. of M and let $M(\mathfrak{p})$ again be the $\mathfrak{p}$-closure of M, i. e., the isolated component of M whose a. p.'s are contained in $\mathfrak{p}$. Let $M'(\mathfrak{p})$ be the isolated component of M whose a. p.'s are *properly* contained in $\mathfrak{p}$. Then, if Q is any $\mathfrak{p}$-primary component of M, $M(\mathfrak{p}) = M'(\mathfrak{p}) \frown Q$. The following lemma then follows immediately from Lemmas **2.** 12 and **2.** 21.

LEMMA **2.** 51. *$M(\mathfrak{p}) \subset M'(\mathfrak{p})$ and $M(\mathfrak{p})$ is $\mathfrak{p}$-primary in $M'(\mathfrak{p})$, namely it is the isolated $\mathfrak{p}$-primary component of M as a sub-module of $M'(\mathfrak{p})$. The $\mathfrak{p}$-primary fundamental ideal of $M(\mathfrak{p})$ as a sub-module of $M'(\mathfrak{p})$ is $M(\mathfrak{p}) : M'(\mathfrak{p}) = Q : M'(\mathfrak{p})$.*

DEFINITION **2.** 51. *The $\mathfrak{p}$-primary ideal $\mathfrak{e}(\mathfrak{p}) = M(\mathfrak{p}) : M'(\mathfrak{p})$ is called the $\mathfrak{p}$-elementary divisor of M. The exponent $\rho(\mathfrak{p})$ of $\mathfrak{e}(\mathfrak{p})$ is called the $\mathfrak{p}$-exponent of M. The length $l(\mathfrak{p})$ of $M(\mathfrak{p})$ as a primary sub-module of $M'(\mathfrak{p})$ is called the $\mathfrak{p}$-length of M.*

Remark **2.** 51. The $\mathfrak{p}$-elementary divisors for $\mathfrak{p} \neq (\omega)$ are the analogues of the primary factors of the elementary divisor ϵ of Part I. Let M be a

module of the vector space of Part I and let $M = \mathrm{Cl}(M) \cap Q_1 \cap \cdots \cap Q_h$ be an *N. D.* of M where Q_j is $\mathfrak{p}_j$-primary and $\mathfrak{p}_j \neq (\omega)$. Then, $M(\mathfrak{p}_j) = \mathrm{Cl}(M) \cap Q_j$, $M'(\mathfrak{p}_j) = \mathrm{Cl}(M)$, and hence $\mathfrak{c}(\mathfrak{p}_j) = Q_j : \mathrm{Cl}(M)$. Since all $\mathfrak{p}_j$'s are relatively prime for $j = 1, \cdots, h$, we conclude that

$$\Pi^h_{j=1}\mathfrak{c}(\mathfrak{p}_j) = \bigcap{}^h_{j=1}\mathfrak{c}(\mathfrak{p}_j) = \bigcap{}^h_{j=1}(Q_j : \mathrm{Cl}(M)) = (\bigcap{}^h_{j=1}Q_j) : \mathrm{Cl}(M)$$
$$= (\mathrm{Cl}(M) \cap \bigcap{}^h_{j=1}Q_j) : \mathrm{Cl}(M) = M : \mathrm{Cl}(M) = \mathfrak{e},$$

which proves the assertion. The interpretation of $\rho(\mathfrak{p}_j)$, given in Theorem **1.12**, is in the general case represented by the fact that $\rho(\mathfrak{p}_j)$ is the exponent of $M(\mathfrak{p}_j) : M'(\mathfrak{p}_j)$, i. e., that $\mathfrak{p}_j^{\rho_j+h}M'(\mathfrak{p}_j) \subseteq M(\mathfrak{p}_j)$ for $h \geq 0$ and not for $h < 0$. Furthermore, $l(\mathfrak{p}_j)$ is the length of $\mathrm{Cl}(M) \cap Q_j$ as a primary sub-module of $\mathrm{Cl}(M)$ from which it follows easily that $l(\mathfrak{p}_j)$ is equal to the multiplicity l_j of the norm $\Delta = \Pi^h_{j=1}\pi^{l_j}{}_j$ of M (see Definition **1.13**). Finally, $M'(\mathfrak{p}_j) = \mathrm{Cl}(M)$ makes it clear that, for a polynomial matrix in n variables, the *different* closures $M'(\mathfrak{p})$ for the a. p.'s of the column space (or for other applications of the row space) are the analogues of the one closure $\mathrm{Cl}(M)$ of a polynomial matrix in one variable.

2.6. Systems of linear equations. The following criterion is the abstract formulation of the known criteria for the solvability of linear equations over different types of scalar domains $\mathfrak{s}$ (see the **Introduction**).

Criterion of Lengths. Let M_1 and M_2 be two modules where $M_1 \subseteq M_2$. Then, $M_1 = M_2$ if and only if M_1 and M_2 have the same a. p.'s and for each a. p. $\mathfrak{p}$, the same $\mathfrak{p}$-length.

Proof. It is obvious that if $M_1 = M_2$, the a. p.'s and $\mathfrak{p}$-lengths of M_1 and M_2 are the same. Hence, we assume that $M_1 \subseteq M_2$ and that M_1 and M_2 have the same a. p.'s and $\mathfrak{p}$-lengths. Let $\mathfrak{p}$ be an a. p. of M_1 and M_2 such that $M_1(\mathfrak{p})$ and hence $M_2(\mathfrak{p})$ have only one a. p., i. e., let $\mathfrak{p}$ be an isolated a. p. of M_1 and M_2. We then conclude from **2.3** that $M_1(\mathfrak{p}) \subseteq M_2(\mathfrak{p})$ and from Lemma **2.51** that $M_1(\mathfrak{p})$ and $M_2(\mathfrak{p})$ are both $\mathfrak{p}$-primary in $M'(\mathfrak{p})$ and from Definition **2.51** that they have the same length as $\mathfrak{p}$-primary sub-modules of $M'(\mathfrak{p})$. Hence $M_1(\mathfrak{p}) = M_2(\mathfrak{p})$ since, otherwise, it would follow immediately from Theorem **2.41** that the length of $M_1(\mathfrak{p})$ is greater than the length of $M_2(\mathfrak{p})$. We then make the induction hypothesis that $M_1(\mathfrak{p}) = M_2(\mathfrak{p})$ has been proved for all a. p.'s for which $M_1(\mathfrak{p})$ and hence $M_2(\mathfrak{p})$ have at most $\kappa - 1$ a. p.'s. Let $\mathfrak{p}$ then be such that $M_1(\mathfrak{p})$ and $M_2(\mathfrak{p})$ have κ a. p.'s. Since $M'_1(\mathfrak{p})$ and $M'_2(\mathfrak{p})$ have at most $\kappa - 1$ a. p.'s, it follows easily from the

induction hypothesis that they are equal, say $M' = M'_1(\mathfrak{p}) = M'_2(\mathfrak{p})$. Then, as before, $M_1(\mathfrak{p})$ and $M_2(\mathfrak{p})$ are $\mathfrak{p}$-primary in M' and have the same length as $\mathfrak{p}$-primary sub-modules of M'; and $M_1(\mathfrak{p}) \subseteq M_2(\mathfrak{p})$. Hence $M_1(\mathfrak{p}) = M_2(\mathfrak{p})$ for all a. p.'s of M, which proves $M_1 = M_2$.

The connection between the criterion of lengths and systems of linear equations is established by considering the m-dimensional Noetherian vector space over $\mathfrak{s}$ which consists of the column vectors whose m components are elements of $\mathfrak{s}$. The columns of an $m \times s$ matrix $A = (\alpha_{ij})$, where $\alpha_{ij} \in \mathfrak{s}$ and $i = 1, \cdots, m$ and $j = 1, \cdots, s$, generate a module M, called the column space of A; and the $\mathfrak{e}(\mathfrak{p})$, $\rho(\mathfrak{p})$ and $l(\mathfrak{p})$ of A are defined as those of the column space.

Criterion of Solvability. Let $\Sigma^s_{j=1}\alpha_{ij}z_j = \gamma_i$, *where* $i = 1, \cdots, m$ *and* α_{ij} *and* γ_i *are elements of* $\mathfrak{s}$, *be a system of* m *linear equations for the* s *unknowns* z_j. *Then, this system has a solution* $z_j \in \mathfrak{s}$ $(j = 1, \cdots, s)$ *if and only if the* $m \times s$ *matrix* $A = (\alpha_{ij})$ *and the augmented* $m \times (s+1)$ *matrix* $B = (\alpha_{ij}, \gamma_i)$ *have the same a. p.'s and for each a. p.* $\mathfrak{p}$, *the same* $\mathfrak{p}$*-length.*

Proof. Let M be the column space of A and g the column vector whose components are $\gamma_1, \cdots, \gamma_m$. The system has a solution if and only if $\gamma \in M$, i e., if and only if $M = L$, where L is the module (M, γ). Since clearly $M \subseteq L$ and L is the column space of B, the criterion of solvability is an immediate corollary of the criterion of lengths.

2. 7. Rank and Cl (M) of a module. In this section, except in Lemma **2.** 71, V will denote a Noetherian vector space over $\mathfrak{s}$ as scalar domain whose zero-module (0) is primary. We want to show that for such a module a notion of rank and Cl (M) exists which has the same properties as in **1. 2** and which is contained in the more general notion of length of a primary composition sequence.

The radical of the zero-module will be denoted by $\mathfrak{r}$; hence $\mathfrak{r}$ is a prime ideal. The radical of an arbitrary module $M \neq (0)$ will be denoted by $\mathfrak{r}(M)$.

DEFINITION **2.** 71. *The closure* Cl (M) *of a module* M *is the* $\mathfrak{r}$*-closure of* M. *To say that* M *is closed or dense means that* M *is respectively* $\mathfrak{r}$*-closed or* $\mathfrak{r}$*-dense.*

It follows from the general theory of **2. 3** that Cl (M) has the properties which are described in **1. 2** and in [6], Section 3. Since $\mathfrak{r} \subseteq \mathfrak{r}(M)$, either M is dense, i. e., Cl (M) $= V$, or Cl (M) is an isolated primary component of

M whose a. p. is $\mathfrak{r}$. In particular, a primary module is either closed or dense (compare with [6], Theorem 4. 2).

Since (0) is $\mathfrak{r}$-primary as a sub-module of V, the zero-module (0) is also $\mathfrak{r}$-primary in any module M (see Lemma **2.** 12).

DEFINITION **2.** 72. *The rank r of a module M is the length of* (0) *as an $\mathfrak{r}$-primary sub-module of M.*

The fact that rank and Cl (M) behave as in **1. 2** follows from the following lemma.

LEMMA **2.** 71. *Let V be an arbitrary Noetherian module with arbitrary zero-module and let Q be a $\mathfrak{p}$-primary module of V. Then, if $Q \subseteq M$, the $\mathfrak{p}$-closure $M(\mathfrak{p})$ is the largest module of V which satisfies the two conditions*: 1) $M \subseteq M(\mathfrak{p})$; 2) *the length of Q as a $\mathfrak{p}$-primary sub-module of M is equal to the length of Q as a $\mathfrak{p}$-primary sub-module of $M(\mathfrak{p})$.*

Proof. The lemma follows immediately from the fact that, in the notation of **2. 3**, $M(\mathfrak{p})$ is the largest module containing M such that $T(M) = T(M(\mathfrak{p}))$.

If we apply Lemma **2.** 71 to the case that Q is the $\mathfrak{r}$-primary zero-module of our Noetherian vector space V, all statements of **1. 2** about rank and Cl (M) are again seen to hold. In particular, if $M_1 \subseteq M_2$ and if M_1 and M_2 have the same rank, then Cl (M_1) = Cl (M_2). We can now reformulate the criterion of lengths, using r instead of the $\mathfrak{r}$-length of M.

THEOREM **2.** 71. *Let V be a Noetherian vector space whose zero-module is primary and let $M_1 \subseteq M_2$ be two modules of V. Then, $M_1 = M_2$, if and only if M_1 and M_2 have the same rank, the same a. p.'s different from $\mathfrak{r}$ and for each such a. p. $\mathfrak{p}$, the same $\mathfrak{p}$-length.*

Proof. The proof follows immediately from the fact that, if $M_1 \subseteq M_2$ and M_1 and M_2 have the same rank, then Cl (M_1) = Cl (M_2). Hence, either M_1 and M_2 do not have $\mathfrak{r}$ as an a. p. and Theorem **2.** 71 immediately reduces to the criterion of lengths; or $\mathfrak{r}$ is an a. p. of M_1 and M_2, in which case M_1 and M_2 have the same isolated $\mathfrak{r}$-primary component Cl (M_1) = Cl (M_2) and hence also the same $\mathfrak{r}$-length.

In order to reformulate also the criterion of solvability, let V be the m-dimensional Noetherian vector space consisting of the column vectors whose m components are elements of $\mathfrak{s}$. We again assume that the zero-module of V is $\mathfrak{r}$-primary and define the rank of a matrix A as the rank of its column

space. The following theorem is then an immediate corollary of the criterion of solvability.

THEOREM **2.72**. *Let* $\Sigma_{j=1}^{s}\alpha_{ij}z_j = \gamma_i$, *where* $i = 1, \cdots, m$ *and* α_{ij} *and* γ_i *are elements of* $\mathfrak{S}$, *be a system of* m *linear equations for the* s *unknowns* z_j. *Then, this system has a solution* $z_j \in \mathfrak{S}$ $(j = 1, \cdots, s)$ *if and only if the* $m \times s$ *matrix* $A = (\alpha_{ij})$ *and the augmented* $m \times (s+1)$ *matrix* $B = (\alpha_{ij}, \gamma_i)$ *have the same rank, the same a. p.'s different from* $\mathfrak{r}$ *and for each such a. p. the same* $\mathfrak{p}$*-length.*

Remark **2.71**. In the vector space of Part I, the zero-module is an (ω)-prime module and hence Theorem **2.72** applies. Since the a. p.'s of a module M which are different from (ω) are then the irreducible factors of the norm $\Delta = \Pi_{j=1}^{h}\pi_j^{l_j}$ of M and since the multiplicity l_j is the (π_j)-length of M, Theorem **2.72** contains Theorem **1.32** as a special case. Since the theory of **1.1**, **1.2** and **1.3** could have been developed just as easily for an arbitrary principal ideal ring as scalar domain as for $P[x]$ (except for the statements about finite P-rank), Theorem **1.32** also holds for principal ideal rings. Actually, the only fact essential for the development of **1.1**, **1.2** and **1.3** is that every factor ring $\mathfrak{S}/\mathfrak{b}$, where $\mathfrak{b}$ is a non-zero ideal of $\mathfrak{S}$, is a principal ideal ring. Hence, Theorem **1.32** also holds if $\mathfrak{S}$ consists of the integers of an algebraic number field (see [4], p. 340). Consequently, the criteria for the solvability of systems of linear equations over principal ideal rings and algebraic number fields are also corollaries of Theorem **2.72**. Finally, if $\mathfrak{S}$ is a field, the zero-module is an (ω)-prime module and every module M is closed. Hence no a. p.'s different from (ω) occur. Theorem **2.72** says that then a system of linear equations can be solved if and only if the matrix A of the system and the augmented matrix B (see Theorem **2.72**) have the same rank. All known criteria of solvability are hence corollaries of Theorem **2.72**, and consequently of the criterion of lengths.

2.8. Rank and linear dependence. In **2.7** the notion of rank is defined independently of the notion of linear dependence. In this section we show that for those vector spaces for which it is customary to define the rank in terms of linear dependence, the notions coincide. This fact follows from Lemma **2.81**.

LEMMA **2.81**. *Let* Q *be a* $\mathfrak{p}$*-prime module of length* l *of a Noetherian vector space* V *over* $\mathfrak{S}$ *as scalar domain. Then,* l *is the maximum number of linearly independent elements of the factor module* V/Q, *if* $\mathfrak{S}/\mathfrak{p}$ *is considered as the scalar domain of* V/Q.

Proof. Since $\mathfrak{p}$ is the fundamental ideal of Q, $\mathfrak{p}$ is the annihilating ideal of V/Q and hence $\mathfrak{s}/\mathfrak{p}$ can be considered as the scalar domain of V/Q. Let $Q \subset Q_1 \subset \cdots \subset Q_{l-1} \subset V = Q_l$ be a $\mathfrak{p}$-primary composition sequence of Q and let $v_j \in Q_j$ where $v_j \notin Q_{j-1}$ for $1 \leq j \leq l$. Each vector v_j belongs to a coset $\bar{v}_j$ of V/Q and we assert that $\bar{v}_1, \cdots, \bar{v}_l$ are linearly independent with respect to $\mathfrak{s}/\mathfrak{p}$. Suppose that $\Sigma^l{}_{j=1}\bar{\alpha}_j\bar{v}_j = \bar{0}$ where $\bar{\alpha}_j \in \mathfrak{s}/\mathfrak{p}$ and $\bar{0}$ is the zero element of V/Q. If α_j is a scalar of the coset $\bar{\alpha}_j$ this would imply that $\Sigma^l{}_{j=1}\alpha_j v_j \in (Q, \mathfrak{p}V)$ and since $\mathfrak{p}V \subseteq Q$, that $\alpha_l v_l \in Q_{l-1}$. However, $v_l \notin Q_{l-1}$ and Q_{l-1} is $\mathfrak{p}$-prime which implies that $\alpha_l \in \mathfrak{p}$ and hence that $\bar{\alpha}_l = \bar{\omega}$ where $\bar{\omega}$ is the zero element of $\mathfrak{s}/\mathfrak{p}$. By repeating this argument for $\Sigma^{l-1}_{j=1}\bar{\alpha}_j\bar{v}_j = \bar{0}$, we find that $\bar{\alpha}_1 = \cdots = \bar{\alpha}_l = \bar{\omega}$ which proves the assertion. There remains to be proved that, if $\bar{w} \in Q/V$, we can find an $\alpha \neq \bar{\omega}$ such that $\Sigma^l{}_{j=1}\bar{\alpha}_j\bar{v}_j = \bar{\alpha}\bar{w}$. Hereto, we assert that $(Q, v_1, \cdots, v_j)(\mathfrak{p}) = Q_j$ for $1 \leq j \leq l$, for $Q \subset (Q, v_1)(\mathfrak{p}) \subseteq Q_1$ and, since $(Q, v_1)(\mathfrak{p})$ is $\mathfrak{p}$-primary, we conclude that $(Q, v_1)(\mathfrak{p}) = Q_1$ from which the assertion follows by induction. In particular, $(Q, v_1, \cdots, v_l)(\mathfrak{p}) = V$ and hence, if w is a vector in the coset $\bar{w}$, we can find an $\alpha \notin \mathfrak{p}$ such that $\alpha w = q + \Sigma^l{}_{j=1}\alpha_j v_j$ where $q \in Q$. This implies that $\bar{\alpha}\bar{w} = \Sigma^l{}_{j=1}\bar{\alpha}_j\bar{v}_j$ where $\bar{\alpha} \neq \bar{\omega}$.

The vector spaces for which it is customary to define rank in terms of linear dependence are m-dimensional vector spaces consisting of columns whose m components are elements of a domain of integrity $\mathfrak{s}$. Hence, for these spaces, the zero-module is a prime module with (ω) as radical. Consequently, the rank of a module M, as defined in Definition **2.** 72, is then the maximal number of vectors of that module which are linearly independent with respect to $\mathfrak{s}$, as follows from Lemma **2.** 81 if we take for Q the (ω)-prime zero-module (0) and as imbedding vector space of (0) the module M.

Remark **2.** 81. If the scalar domain of a Noetherian vector space has divisors of zero, the notion of linear dependence does not give rise to useful notions of rank . However, Definition **2.** 72 defines rank with all the desirable useful properties for any scalar domain $\mathfrak{s}$ (with or without divisors of zero) for the case when the zero-module is primary. Moreover, it is never really necessary to make any definition of rank since, according to **2. 7**, a definition of rank is nothing more than a preferential treatment of that primary component of a module whose a. p. is the radical of the zero-module.

PRINCETON UNIVERSITY,
PRINCETON, NEW JERSEY.

REFERENCES.

[1] R. A. Frazer, W. J. Duncan and A. R. Collar, *Elementary Matrices*, Cambridge University press (1938).

[2] A. A. Albert, *Modern Higher Algebra*, University of Chicago Press (1936).

[3] K. Hentzelt and E. Noether, "Zur Theorie der Polynomideale und Resultanten," *Mathematische Annalen*, vol. 88 (1922), pp. 53-79.

[4] E. Steinitz, "Rechteckige Systeme und Moduln in algebraischen Zahlkorpern I," *Mathematische Annalen*, vol. 71 (1912), p. 340.

[5] P. M. Grundy, "A generalization of additive ideal theory," *Proceedings of the Cambridge Philosophical Society*, vol. 38 (1941), pp. 241-279.

[6] E. Snapper, "Structure of linear sets," *Transactions of the American Mathematical Society*, vol. 52 (1942), pp. 257-264.

[7] W. Krull, "Idealtheorie," *Ergebnisse der Mathematik*, vol. 4 (1935), pp. 255-402.

[8] B. L. v. d. Waerden, "Eine Verallgemeinerung des Bezoutschen Theorems," *Mathematische Annalen*, vol. 99 (1928), pp. 497-541.

[9] B. L. v. d. Waerden, *Moderne Algebra*, vol. 1, Berlin, Springer, 1937.

[10] B. L. v. d. Waerden, "On Hilbert's function, series of composition of ideals and a generalization of the theorem of Bezout," *Akademie van Wetenschappen te Amsterdam*, Proc., vol. 31, 1928, pp. 749-770.

[11] B. L. v. d. Waerden, *Moderne Algebra*, vol. 2, Berlin, Springer, 1940.

[12] C. Chevalley, "On the theory of local rings," *Annals of Mathematics*, vol. 44 (1943), pp. 690-708.

[13] E. Snapper, "The resultant of a linear set," *American Journal of Mathematics*, vol. 66 (1944), pp. 59-68.

[14] C. Chevalley, "On the notion of the ring of quotients of a prime ideal," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 93-97.

SYSTEMS OF TOTAL DIFFERENTIAL EQUATIONS AND LIOUVILLE'S THEOREM ON CONFORMAL MAPPINGS.*

By PHILIP HARTMAN.

1. Let X denote the vector (x_1, x_2, x_3) and let the vector function $X = X(u, v)$, defined on a connected domain R in the (u, v)-plane, be a surface S of class C''; so that each component of $X(u, v)$ has continuous partial derivatives of the second order and the vector product

$$(X_u, X_v) \neq 0 \text{ on } R, \tag{1}$$

where the subscript u or v indicates the corresponding partial derivative. Since the surface S is of class C'', every normal section of S is a curve of class C'', and so possesses a curvature. At the point $X(u, v)$, let $k_1(u, v)$ and $k_2(u, v)$ denote the principal curvatures, the maximum and minimum curvatures of all normal sections at the point $X(u, v)$. It is known that if every point of S is a "flat point" ("Flachpunkt"), $k_1(u, v) = k_2(u, v) \equiv 0$ on R, then the surface S is part of a plane. Also, if every point of the surface is an "umbilicus" ("Nabelpunkt"), $k_1(u, v) = k_2(u, v) \neq 0$ on R, and if, *in addition,* it is supposed that $X(u, v)$ is of class C''', then it is known that the surface S is part of a sphere. Cf., e. g., Bieberbach [2].

It will be shown below that the *additional* assumption of the last statement, requiring that $X(u, v)$ be of class C''', rather than C'', is not needed. It can hardly be expected to relax the C''-condition since the definition of the principal curvatures involves the second fundamental form of the surface. In this direction the following theorem will be proved:

THEOREM 1. *Let $X = X(u, v)$, defined on a connected domain R, be a surface S of class C''. Let $k_1(u, v)$ and $k_2(u, v)$ denote the principal curvatures and let*

$$k_1(u, v) \equiv k_2(u, v) \text{ on } R. \tag{2}$$

Then the function (2) is constant, and so the surface S is part of either a plane or a sphere, according as this constant is or is not zero.

* Received October 18, 1946.

2. The proof will depend on a theorem on total differentials which may be considered an analogue of the lemma of P. du Bois-Reymond occurring in the calculus of variations. This theorem can be formulated as follows:

THEOREM 2. *Let*

$$x = x(u, v), \qquad y = y(u, v) \tag{3}$$

denote scalar functions of class C' on a connected domain U of the (u, v)-plane and suppose that the Jacobian

$$x_u y_v - x_v y_u \neq 0 \text{ on } U. \tag{4}$$

Let $\alpha(u, v)$ be a continuous function on U. A necessary and sufficient condition for

$$\int_C \alpha(u, v)\, dx(u, v) = 0 \text{ and } \int_C \alpha(u, v)\, dy(u, v) = 0 \tag{5}$$

to hold for every rectifiable Jordan curve C in U is that $\alpha(u, v)$ be a constant.

If the functions $x(u, v)$ and $y(u, v)$ possess continuous partial derivatives of the second order and if the function $\alpha(u, v)$ possesses partial derivatives of the first order, then necessary and sufficient conditions for (5) are

$$(\alpha x_u)_v = (\alpha x_v)_u \text{ and } (\alpha y_u)_v = (\alpha y_v)_u.$$

Performing the partial differentiations indicated in these equations and using the condition (4), one obtains the identities $\alpha_u \equiv \alpha_v \equiv 0$. However, the point in the theorem lies in the fact that these operations are not justifiable under the conditions of Theorem 2, since the existence of the partial derivatives involved is not assumed.

The necessity of the condition (5) is obvious.

Let (u_0, v_0) be an arbitrary point of the domain U. In order to prove the sufficiency of (5), it will be shown that if (5) holds for every rectifiable Jordan curve C in U, then there is a neighborhood of (u_0, v_0) on which $\alpha(u, v)$ is constant.

Let $x_0 = x(u_0, v_0)$ and $y_0 = y(u_0, v_0)$. Then (4) implies that (3) possesses a unique local inverse,

$$u = u(x, y), \qquad v = v(x, y), \tag{6}$$

of class C' which transforms some neighborhood,

$$D: |x - x_0| < d, \qquad |y - y_0| < d \qquad (d > 0), \tag{7}$$

of (x_0, y_0) into a neighborhood of (u_0, v_0). Thus, it will be sufficient to show that the continuous function

$$\beta(x, y) = \alpha(u(x, y), v(x, y)), \tag{8}$$

defined on (7), is constant.

Let the curve $C: u = u(t)$, $v = v(t)$, be the image, under the transformation (6), of the rectangle whose vertices are (x_0, y_0), (x, y_0), (x, y), (x_0, y), where (x, y) is any point of D and the parameter t denotes arc-length on the rectangle (in the (x, y)-plane). Since (6) is of class C', the Jordan curve C is rectifiable. Thus, the first equation in (5) gives

$$\int_{x_0}^{x} \beta(t, y_0)\, dt - \int_{x_0}^{x} \beta(t, y)\, dt = 0.$$

Since the function (8) is continuous, it follows from the fundamental theorem of calculus that

$$\beta(x, y_0) = \beta(x, y),$$

if (x, y_0) and (x, y) are points of the domain D. Similarly, the second equation of (5) implies that

$$\beta(x_0, y) = \beta(x, y),$$

if (x_0, y) and (x, y) are points of (7). The last two formulae mean that $\beta(x, y)$ is constant on (7). This proves Theorem 2.

3. In order to prove Theorem 1, use will be made of the differential equations of Olinde Rodrigues for the lines of curvature on a surface of class C'',

$$N' + \alpha X' = 0, \tag{9}$$

where $N = N(u, v)$ is the unit normal vector, $(X_u, X_v)/|(X_u, X_v)|$; the function $\alpha = \alpha(u, v)$ is either $k_1(u, v)$ or $k_2(u, v)$; and the prime denotes differentiation with respect to a parameter which does not occur explicitly in the differential equations (9).

Under the assumption (2) of Theorem 2, every arc of class C'' on the surface is a line of curvature, that is, a solution of (9). Hence,

$$\int_C \alpha(u, v)\, dX(u, v) = -\int_C dN(u, v) = 0 \tag{10}$$

holds for every Jordan curve C of class C'' on R, the domain of definition of the surface $X(u, v)$. By the 2-dimensional analogue of Helly's term-by-term integration theorems for Stieltjes integrals and in view of standard approximation theorems of the Weierstrass type, (10) holds also for every rectifiable Jordan curve C on R. Since the components of the vector in (1) are the three Jacobians

$$x_{2u}x_{3v} - x_{2v}x_{3u}; \qquad x_{3u}x_{1v} - x_{3v}x_{1u}; \qquad x_{1u}x_{2v} - x_{1v}x_{2u},$$

it follows from (1) that, at every point of R, at least one of these Jacobians is not zero. Let P be an arbitrary point of R and let x, y denote a pair of the variables x_1, x_2, x_3 selected so that $x_u y_v - x_v y_u \neq 0$ at P. Then, by continuity, (4) holds on some neighborhood U of P. On the other hand, the vector equation (10) implies that (5) holds for every rectifiable Jordan curve C on U. Consequently, by Theorem 2, the continuous function $\alpha(u, v)$ is constant on U. Since the point P of the connected domain R was arbitrary, it follows that the function $\alpha(u, v)$ or (2) is constant on R.

This completes the proof of the first statement of Theorem 1. The second statement is an immediate consequence of the first. For, since $\alpha = \alpha(u, v)$ is a constant, an integration of (9) gives

$$N(u, v) + \alpha X(u, v) = A,$$

a constant vector. If $\alpha = 0$, then the normal vector $N = N(u, v)$ is this constant vector A. But along any arc (of class C') on the surface, the tangent vector is orthogonal to $N = A$. This means that the scalar product $A \cdot X'$ vanishes for all t, so that $A \cdot X = c$, where c is a constant. This shows that the surface S is part of a plane. Similarly, if $\alpha \neq 0$, then

$$X(u, v) = (A - N(u, v))/\alpha,$$

and so, since $N(u, v)$ is a unit vector, the surface S lies on the sphere having the center A/α and radius $|1/\alpha|$.

This completes the proof of Theorem 1.

4. Let

$$Y = Y(X) \tag{11}$$

where $X = (x_1, x_2, x_3)$, $Y = (y_1, y_2, y_3)$ and

$$\det(\partial y_i/\partial x_k) \neq 0$$

be a (schlicht) mapping of class C' of a domain in the X-space upon a domain in the Y-space. The transformation (11) is said to be conformal if there exists a positive, continuous function $\lambda = \lambda(X)$ on T such that

$$dy_1^2 + dy_2^2 + dy_3^2 = \lambda^2(dx_1^2 + dx_2^2 + dx_3^2);$$

with a corresponding definition of comformality in the case of a space with an arbitrary dimension number. A mapping (11) is said to be a Möbius transformation if it can be composed of Euclidean movements, reflections on spheres or planes, and similarity transformations ($Y = cX$, where c is a scalar constant).

Liouville's theorem states that (under conditions of differentiability not specified by him) every conformal mapping is a Möbius transformation if the dimension number of the space exceeds 2. It was pointed out to me by Professor Wintner that the various proofs of Liouville's theorem as given in texts on differential geometry assume either explicitly or implicitly that the conformal mapping (11) is of class C'''. He raised the question whether the class C''' can be reduced to the class C' (just as in the plane case of the Cauchy-Riemann equations, where the Möbius group must be replaced by the group of regular analytic schlicht mappings).

Although the above results do not apply to this extent, they halve the gap between the desideratum and the standard condition, since they reduce the class C''' to the class C''.

THEOREM 3. *Every 3-dimensional conformal mapping of class C'' is a Möbius transformation.*

A geometrical proof of Liouville's theorem given by Cappelli [4] and reproduced by Blaschke [3] (without bothering with the assumptions to be made) is based on

(i) Dupin's theorem that if three surfaces of class C'' cut orthogonally, then the arcs of intersection are lines of curvature on each surface, and on

(ii) the circumstance that, if every direction at every point on the surface is a direction of a line of curvature, then the surface is part of either a plane or a sphere.

Since, as proved above, (ii) holds for surfaces of class C'', Capelli's proof becomes valid for the sharpened formulation of Liouville's assertion, that is, for Theorem 3.

Remark. Another proof of Liouville's theorem, as given by Bianchi [1], replaces the geometrical considerations (i) and (ii) by Lamé's partial differential equations for orthogonal families. This proof, too, assumes the mapping (11) to be of class C'''. However, due to Theorem 2, the derivation of Lamé's differential equations can be modified so as to lead to another proof of Theorem 3.

The methods of this proof are applicable to the case of any n-dimensional ($n \geqq 3$) space. Actually, the previous proof based on (i) and (ii) can be modified to this effect, since Theorems 1 and 2 have n-dimensional analogues.

THE JOHNS HOPKINS UNIVERSITY.

BIBLIOGRAPHY.

[1] L. Bianchi, *Vorlesungen über Differentialgeometrie*, Leipzig, 1899, pp. 484-488.
[2] L. Bieberbach, *Differentialgeometrie*, Leipzig and Berlin, 1932, pp. 68-69.
[3] W. Blaschke, *Vorlesungen über Differentialgeometrie*, vol. I, Berlin, 1930, pp. 100-102.
[4] A. Capelli, "Sulla limitata possibilità di trasformazioni conformi nello spazio," *Annali di Matematica*, Ser. II, vol. 14 (1886-1887), pp. 227-237.

GENERIC ALGEBRAS.* [1]

By W. J. R. Crosby.

The aim of this paper is to establish a relation between a given algebra class over a suitable type of ground field and certain other algebras, called generic algebras. Suppose that P is the ground field of the given algebra class; then types of central simple algebras are constructed over algebraic function fields with P as field of constants. These algebras are called generic algebras. It is shown that the original algebra class may be obtained from the generic algebra by a specialization of the variables. Thus problems in the theory of algebras may be associated with corresponding problems in the theory of algebraic functions of several variables. In particular it is seen that, in order to consider certain questions on algebras such as that of the possibility of the representation of an algebra as a direct product or crossed product or the existence of cyclic splitting fields of minimal degree, it is sufficient to confine our attention to generic algebras.

After a review in **1** of the theory of algebras needed subsequently, **2** contains a definition of suitable fields of algebraic functions associated with given algebra classes and a definition of the class of related generic algebras. In **3** it is shown how the generic algebra specializes to an algebra of the given algebra class, and in **4** some of the properties of the generic algebras are discussed. Finally, in **5**, it is shown that the properties of a special type of generic algebras, called regular generic algebras, may be more easily discussed.

Generic Algebras.

1. Suppose that S_0 is a central [2] simple algebra over the field P. Then there exists an algebraic extension F of P, which is normal and separable over P, and which is a splitting field [3] for S_0; that is

* Received May 18, 1945.

[1] This dissertation was presented to the University of Toronto as part of the requirements for the degree of Doctor of Philosophy. The author wishes to record his gratitude to Professor R. Brauer, under whose supervision the paper was written, and to Dean S. Beatty.

[2] Or normal.

[3] See M. Deuring: Algebren, *Ergebnisse der Mathematik*, vol. 4 (1935), Berlin, p. 48. Also N. Jacobson: *The Theory of Rings*, Mathematical Surveys 2, American Mathematical Society, New York, 1943, p. 104, and B. L. van der Waerden, *Moderne Algebra*, vol. 2 (1940), Berlin, p. 209.

$$S_0 \times F \cong [F]_{r_0},$$

where $[F]_{r_0}$ is the complete matrix algebra of degree r_0 with coefficients in F: The order of S_0 over P is the number r_0^2. Also, by Wedderburn's theorem,[4] S_0 is isomorphic to the complete algebra of matrices of some degree t_0, with coefficients in a division ring D, and this division ring D is unique in the sense of isomorphism; that is

$$S_0 \cong [D]_{t_0} \cong D \times [P]_{t_0}.$$

The division ring D, hereafter referred to as the Wedderburn factor of S_0, is itself a central (and simple) algebra with center P, and if m^2 is its order over P, then $r_0 = t_0 m$. The unique number m so determined is the index of S_0.

In the splitting field F there exists an absolutely irreducible representation of the algebra S_0 by means of matrices of order r_0. The absolutely irreducible representation and its conjugates determine a factor set c of S_0, the elements of which we shall denote by $c_{A,B,C}$, where A,B,C are elements of the Galois group $\mathfrak{G}$ of F over P. About these numbers $c_{A,B,C}$ we can make the following[5] statements:

1) $c_{A,B,C} \neq 0$, for all A,B,C of $\mathfrak{G}$.

2) $c^G{}_{A,B,C} = c_{AG,BG,CG}$, where c^G denotes the transform of c by the element G of $\mathfrak{G}$.

3) $c_{A,B,C} \cdot c_{A,C,G} = c_{A,B,G} \cdot c_{B,C,G}$, for all elements A,B,C,G of $\mathfrak{G}$.

4) There is a positive integer e, and there are non-zero numbers $k_{A,B}$ in F, corresponding to each pair of elements A,B of $\mathfrak{G}$, which form a conjugate double system,[6] such that

$$c^e{}_{A,B,C} = \frac{k_{A,B} \cdot k_{B,C}}{k_{A,C}},$$

for all elements A,B,C of $\mathfrak{G}$. The least positive number e with this property is the exponent of the factor set c in the field F.

5) The exponent is a divisor of the index and is divisible by all prime factors of the index.

---

[4] See N. Jacobson, *loc. cit.*, p. 98.

[5] The first three might form a definition of a factor set, the others may be regarded as consequences of the first three and the related theory of algebras. See M. Deuring, *loc. cit.*, p. 52 *et seq.*

[6] That is, the transform of $k_{A,B}$ by the automorphism G of the Galois group $\mathfrak{G}$ is $k_{AG,BG}$.

6) The factor set c is not uniquely determined by S_0 and F, but may be replaced by the set c', the elements of which are defined by

$$c'_{A,B,C} = \frac{k_{A,B} \cdot k_{B,C}}{k_{A,C}} \cdot c_{A,B,C},$$

where $k_{A,B}$ is any conjugate double system with no zero elements. The factor set c' is called an associated factor set.

7) If the Wedderburn factor of an algebra S_1, central over P, is isomorphic to the Wedderburn factor of S_0, then the algebra S_1 determines, with F, the same class of associated factor sets as that determined by S_0.

8) The set of all matrices of the form $(z_{A,B}\, c_{A,B,E})$ where $z_{A,B}$ may be any conjugate double system in F, and E is the unit element of $\mathfrak{G}$, constitute a central simple algebra of which the order is the square of the order r of $\mathfrak{G}$, and of which the Wedderburn factor is D, the Wedderburn factor of S_0.

2. In order to be able to construct suitable algebraic function fields associated with the algebra S, stated to exist in 8), we shall assume[7] that all the elements of the factor set c lie in P.

Let T denote a $(1-1)$ representation of $\mathfrak{G}$ by means of matrices of degree q with coefficients in P. Such a representation always exists. We shall be concerned with extension fields of P, formed by adjoining q algebraic or transcendental elements to K. If $\mathfrak{p}$ is a column of length q, with elements $p_1, p_2, \cdots, p_q$, we shall write $P(\mathfrak{p})$ for $P(p_1, p_2, \cdots, p_q)$; that is, $P(\mathfrak{p})$ is the field of all rational functions, with non-zero denominators, of $p_1, p_2, \cdots, p_q$, with coefficients in P. Similarly, we shall write $f(\mathfrak{p})$ for $f(p_1, p_2, \cdots, p_q)$, where f is a rational function of $p_1, p_2, \cdots, p_q$ (with non-zero denominator). Let $\mathfrak{x}$ be the column of length q with elements $x_1, x_2, \cdots, x_q$, where $x_1, x_2, \cdots, x_q$ are indeterminates over F. Let $\Phi = P(\mathfrak{x})$, and let Π be the field of all elements of Φ which are invariant under the transformations of the representation T of $\mathfrak{G}$ in P. That is, if T_G is the matrix of the representation which corresponds to the element G of $\mathfrak{G}$, then Π is the field of all rational functions $f(\mathfrak{x})$ of Φ which are such that $f(T_G\mathfrak{x}) = f(\mathfrak{x})$ for every G in $\mathfrak{G}$. If we write $\mathfrak{x}^G$ for $T_G\mathfrak{x}$, and $f(\mathfrak{x})^G$, or f^G, for $f(T_G\mathfrak{x})$, then we shall have

$$\begin{aligned}(f(\mathfrak{x})^G)^H = f(T_G\mathfrak{x})^H = f(T_G\mathfrak{x}^H) = f(T_G T_H \mathfrak{x}) \\ = f(T_{GH}\mathfrak{x}) = f(\mathfrak{x}^{GH}) = f^{GH},\end{aligned}$$

as we should expect.

---

[7] We shall see later that we may instead assume that P contains certain roots of unity.

The field Φ is normal[8] over Π. For, if f is any element of Φ, and t is an indeterminate, then evidently the function $\chi(t) = \prod_G (t - f^G)$, where the product is taken over all elements G of $\mathfrak{G}$, is a polynomial with coefficients in Π. All the roots of $\chi(t) = 0$ lie in Φ, and one of the roots is f. Since every element of Φ is the root of an equation of degree less than or equal to r with coefficients in Π, Φ is normal over Π with degree not greater than r. But certainly all the elements of $\mathfrak{G}$ define automorphisms of Φ over Π. Hence the degree of Φ over Π is exactly r. The Galois group of Φ over Π is evidently $\mathfrak{G}$. Also the element G of $\mathfrak{G}$ carries f into f^G, in accordance with the notation used above.

We next define an algebra Σ by means of Φ in the same way that we defined S by means of F. Since the factor set c lies in P, it must lie in Φ. Let $\zeta_{A,B}$ be a conjugate double system of elements of Φ. Let Σ be the set of all matrices of the form $(\zeta_{A,B}\, c_{A,B,E})$. Then Σ will be called the generic algebra associated with S; it is a central simple algebra or order r with center Π, and is determined by the fields P and F and the factor set c. In order to learn in what ways Σ over Π resembles S over P, we try to replace the indeterminates of Φ by elements of F, in such a way that, in this process of specialization, Φ becomes F, Π becomes P, and Σ becomes S.

3. The results on specialization of variables which will be established all depend on the properties of a certain matrix. In our first lemma we prove the existence of this matrix. The letter $\mathfrak{p}$, in all our arguments, will be used to denote a column of length q with elements $p_1, p_2, \cdots, p_q$; the numbers $p_1, p_2, \cdots, p_q$ will always be assumed to be elements of P.

LEMMA 1. *There exists a matrix R of degree q with coefficients in F, such that*

(i) $R^G = T_G R$;

(ii) *R is non-singular;*

(iii) *If the vector $\mathfrak{p}$ is invariant under no transformation T_G except when G is the unit element E of $\mathfrak{G}$, then $R\mathfrak{p}$ is invariant for only the unit element E of $\mathfrak{G}$; that is, if the relation $T_G\mathfrak{p} = \mathfrak{p}$ implies that $G = E$, then the relation $(Rp)^G = Rp$ implies that $G = E$;*

(iv) *If the representation T is irreducible, and $\mathfrak{p}$ is not zero, then $R\mathfrak{p}$ is invariant for only the unit element E of $\mathfrak{G}$.*

---

[8] Cf., for instance, E. Artin, *Galois Theory*, Notre Dame Mathematical Lectures 2, Notre Dame, 1942, p. 31.

Proof. Let v be a number of F such that the transforms of v by the elements of $\mathfrak{G}$ are linearly independent [9] with respect to P. Then the system of numbers $u_G = v^{G^{-1}}$ forms a basis for F over P, and

$$u_G{}^H = v^{G^{-1}H} = v^{(H^{-1}G)^{-1}} = u_{H^{-1}G}$$

for every pair of elements G,H of $\mathfrak{G}$. We now prove that we may take

$$R = \Sigma_G u_G T_G,$$

where the summation extends over all elements of $\mathfrak{G}$.

The statement (i) is certainly true, for, writing L for $G^{-1}H$, we have

$$R^G = \Sigma_H u^G{}_H T_H = \Sigma_H u_{G^{-1}H} T_H$$
$$= \Sigma_L u_L T_{GL} = T_G \Sigma_L u_L T_L = T_G R.$$

Next let $\mathfrak{y}$ be a column of length q with elements $y_1, y_2, \cdots, y_q$, and let us suppose that the equation $R\mathfrak{y} = 0$ has a non-trivial solution. Let $F = P(\theta)$ where θ is an element of F and let us write θ_G for θ^G. Then since $R\mathfrak{y} = 0$, $(R\mathfrak{y})^G = 0$. Therefore $R(\mathfrak{y}^G) = 0$, since $R^G = T_G R$ and T_G is non-singular. Consequently

$$R\Sigma_G \theta_G{}^j \mathfrak{y}^G = \Sigma_G \theta_G{}^j R(\mathfrak{y}^G) = 0$$

for any positive integer j. But $\Sigma_G \theta_G{}^j \mathfrak{y}^G$ is different from zero if j is suitably chosen and all its elements lie in P. We have therefore proved that if the equation $R\mathfrak{y} = 0$ has a non-trivial solution in F, it has a non-trivial solution in P. But if we now suppose that $R\mathfrak{y} = 0$ where $y_1, y_2, \cdots, y_q$ lie in P, then from the definition of R, $\Sigma_G u_G T_G \mathfrak{y} = 0$, and consequently $T_G \mathfrak{y} = 0$ for every G in $\mathfrak{G}$, since the system of numbers u_G forms a basis for F over P, and the elements of T_G are in P. Therefore $\mathfrak{y} = 0$, since T_G is non-singular. The equation $R\mathfrak{y} = 0$ has, therefore, only the trivial solution $\mathfrak{y} = 0$. Consequently R is non-singular. This is part (ii) of the lemma.

Let us now suppose that $(R\mathfrak{p})^G = R\mathfrak{p}$. Employing part (i) of the lemma, we can write this relation in the form $(T_G - I)R\mathfrak{p} = 0$, where I is the unit matrix, and substituting our expression for R, we find

$$\Sigma_H (T_G - I) u_H T_H \mathfrak{p} = 0.$$

From this equation we conclude that

$$(T_G - I) T_H \mathfrak{p} = 0$$

for every H in $\mathfrak{G}$. This implies that $T_{GH}\mathfrak{p} = T_H \mathfrak{p}$ and hence, that

[9] See E. Artin, *loc. cit.*, p. 53.

$$T_{H^{-1}GH}\mathfrak{p} = \mathfrak{p} \tag{1}$$

for every H in $\mathfrak{G}$.

The statement (iii) now follows; for if the relation $T_G\mathfrak{p} = p$ implies that $G = E$, the equation (1) implies that $H^{-1}GH = E$ and therefore $G = E$.

Finally, to prove (iv), we first note that, if $(R\mathfrak{p})^G = R\mathfrak{p}$, then from the equation (1) which we have established we conclude that p is invariant under any transformation T_M, if M is any conjugate element or product of conjugate elements of G. Let $\mathfrak{M}$ be the class of all the conjugates and products of conjugates of G. We easily verify that $\mathfrak{M}$ is an invariant subgroup of $\mathfrak{G}$. Assuming that $\mathfrak{p}$ is not zero, let us take the elements of $\mathfrak{p}$ to be the components of the first of q linearly independent vectors in a q-dimensional vector space over P. Then the set of vectors with only the first component different from zero constitutes an invariant subspace with respect to the linear transformations defined by the representation T_M of the group $\mathfrak{M}$. Therefore, there must be a non-singular matrix Q with coefficients in P, such that

$$Q^{-1}T_MQ = \begin{pmatrix} 1 & 0 \\ 0 & U_M \end{pmatrix}$$

for every M in $\mathfrak{M}$, where U_M is a representation of $\mathfrak{M}$ of degree $q - 1$. Therefore, by Clifford's Theorem,[10] the representation T_M of $\mathfrak{M}$ reduces completely to the unit matrix. Consequently the subgroup $\mathfrak{M}$ must contain only the unit element of $\mathfrak{G}$. Thus $(R\mathfrak{p})^G = R\mathfrak{p}$ implies that G is the unit element of $\mathfrak{G}$.

We can now discuss the result of replacing the indeterminates in $\Phi = P(\mathfrak{x})$ by suitably chosen elements of F. It is, of course, to be understood that elements of which the denominators are made to vanish by the substitution are to be discarded.

THEOREM 1. *If we replace $\mathfrak{x}$ in $\Phi = P(\mathfrak{x})$ by $R\mathfrak{p}$, the resulting specialized field, namely $P(R\mathfrak{p})$, coincides with F, provided either that $\mathfrak{p}$ satisfies no equation of the form $T_G\mathfrak{p} = \mathfrak{p}$ except when G is the unit element E of $\mathfrak{G}$, or that T is irreducible and $\mathfrak{p}$ is not zero.*

Proof. Since $P(R\mathfrak{p})$ contains P and is contained in F, $P(R\mathfrak{p})$ must be the field of invariants for some subgroup $\mathfrak{A}$ of the Galois group $\mathfrak{G}$ of F over P. Then $R\mathfrak{p}$ must be invariant under the transformations of $\mathfrak{A}$. That is, $(R\mathfrak{p})^G = R\mathfrak{p}$ if G is in $\mathfrak{A}$. Therefore, by Lemma 1, $\mathfrak{A}$ contains only the unit element of $\mathfrak{G}$. Therefore $P(R\mathfrak{p}) = F$.

---

[10] See H. Weyl, *The Classical Groups*, Princeton, 1939, p. 159.

THEOREM 2. *If we replace* $\mathfrak{x}$ *in each function* $f(\mathfrak{x})$ *of* Π *by* $R\mathfrak{p}$, *the resulting specialized field coincides with* P.

Proof. Let us denote the specialized field by P_1. Evidently F contains P_1. Since Π contains P, P_1 must contain P. Now Π was defined to be the set of all rational functions $f(\mathfrak{x})$, with coefficients in P, such that $f(T_G\mathfrak{x}) = f(\mathfrak{x})$. If we replace $\mathfrak{x}$ by $R\mathfrak{p}$ in these functions, and discard those expressions of which the denominator vanishes, we find that, for every element $f(R\mathfrak{p})$ of P_1,

$$f(R\mathfrak{p})^G = f(R^G\mathfrak{p}) = f(T_G R\mathfrak{p}) = f(R\mathfrak{p})$$

for every G in $\mathfrak{G}$. It follows that P_1 is contained in P, and therefore $P_1 = P$.

THEOREM 3. *If we replace* $\mathfrak{x}$ *by* $R\mathfrak{p}$ *in* Σ, *the specialized algebra coincides with* S, *provided either that* $\mathfrak{p}$ *satisfies no equation of the form* $T_G\mathfrak{p} = \mathfrak{p}$ *except when* G *is the unit element* E *of* $\mathfrak{G}$, *or that* T *is irreducible and* $\mathfrak{p}$ *is different from zero.*

Remark. It is to be understood that, if the substitution of $R\mathfrak{p}$ for $\mathfrak{x}$ makes the denominator of one of the coefficients of a matrix of Σ vanish, that matrix is to be discarded.

Proof. It is evidently sufficient to prove that (i) if $\zeta_{A,B}$ is a conjugate double system of elements of Φ, and if the system $\zeta_{A,B}$ specializes to the system $z_{A,B}$ in F, when we replace $\mathfrak{x}$ by $R\mathfrak{p}$, then $z_{A,B}$ is a conjugate double system in F, and also that (ii) if $z_{A,B}$ is any conjugate double system of elements of F, there exists in Φ a conjugate double system $\zeta_{A,B}$ which reduces to $z_{A,B}$ when we replace $\mathfrak{x}$ by $R\mathfrak{p}$. The first statement is certainly true, for if $\zeta_{A,B} = \zeta_{A,B}(\mathfrak{x})$, then $z_{A,B} = \zeta_{A,B}(R\mathfrak{p})$, and consequently

$$z^G{}_{A,B} = \zeta_{A,B}(R\mathfrak{p})^G = \zeta_{A,B}(T_G R\mathfrak{p})$$
$$= \zeta_{AG,BG}(R\mathfrak{p}) = z_{AG,BG}.$$

In order to establish (ii), let us suppose that $z_{A,B}$ is a conjugate double system in F. Let E be the unit element of $\mathfrak{G}$. Then by Theorem 1, there are elements $\zeta_{E,A} = \zeta_{E,A}(\mathfrak{x})$ in Φ, which specialize to the elements $z_{E,A}$ in F for each A in $\mathfrak{G}$. We then define $\zeta_{A,B}$ to be $\zeta^A{}_{E,BA^{-1}} = \zeta_{E,BA^{-1}}(T_A\mathfrak{x})$. The system so defined is a conjugate double system in Φ, since

$$\zeta^G{}_{A,B} = \zeta_{E,BA^{-1}}(T_A\mathfrak{x}^G) = \zeta_{E,BA^{-1}}(T_A T_G\mathfrak{x})$$
$$= \zeta_{E,(BG)(AG)^{-1}}(T_{AG}\mathfrak{x}) = \zeta_{AG,BG}.$$

Also, $\zeta_{A,B}$ specializes to $z_{A,B}$ for every A,B of $\mathfrak{G}$, since

$$\zeta_{A,B}(R\mathfrak{p}) = \zeta_{E,BA^{-1}}(T_A R\mathfrak{p}) = \zeta_{E,BA^{-1}}(R^A\mathfrak{p})$$
$$(\zeta_{E,BA^{-1}}(R\mathfrak{p}))^A = z^A{}_{E,BA^{-1}} = z_{A,B}.$$

The proof is now complete.

4. In order to discover the properties of Σ it is necessary to examine a little more closely the conjugate double systems $\zeta_{A,B}$, which are characterized by the relation $\zeta^G{}_{A,B} = \zeta_{AG,BG}$. Suppose $\zeta_{A,B} = \xi_{A,B}(\mathfrak{x})/\eta_{A,B}(\mathfrak{x})$, where $\xi_{A,B}(\mathfrak{x})$ and $\eta_{A,B}(\mathfrak{x})$ are polynomials which are prime to each other. Then since

$$\left(\frac{\xi_{A,B}(\mathfrak{x})}{\eta_{A,B}(\mathfrak{x})}\right)^G = \frac{\xi_{A,B}(T_G\mathfrak{x})}{\eta_{A,B}(T_G\mathfrak{x})} = \frac{\xi_{AG,BG}(\mathfrak{x})}{\eta_{AG,BG}(\mathfrak{x})},$$

we have

$$\xi_{A,B}(T_G\mathfrak{x}) \cdot \eta_{AG,BG}(\mathfrak{x}) = \xi_{AG,BG}(\mathfrak{x}) \cdot \eta_{A,B}(T_G\mathfrak{x}). \tag{1}$$

Therefore $\xi_{AG,BG}(\mathfrak{x})$, being prime to $\eta_{AG,BG}(\mathfrak{x})$, must divide $\xi_{A,B}(T_G\mathfrak{x})$. Consequently

$$\xi^G{}_{A,B}(\mathfrak{x}) = \xi_{A,B}(T_G\mathfrak{x}) = m_{A,B,G}\xi_{AG,BG}(\mathfrak{x}),$$

where $m_{A,B,G}$ is an element of P, since $\xi_{A,B}(T_G\mathfrak{x})$ and $\xi_{AG,BG}(\mathfrak{x})$ have the same degree. From the relation (1) we see that we have also

$$\eta^G{}_{A,B}(\mathfrak{x}) = \eta_{A,B}(T_G\mathfrak{x}) = m_{A,B,G}\eta_{AG,BG}(\mathfrak{x}).$$

Let $m_G = (\Pi_A\Pi_B m_{A,B,G})^2$, where A and B run through the elements of $\mathfrak{G}$. Similarly let $M(\mathfrak{x}) = \Pi_A\Pi_B\xi_{A,B}(\mathfrak{x}) \cdot \eta_{A,B}(\mathfrak{x})$. Then

$$M^G = M(T_G\mathfrak{x}) = m_G M(\mathfrak{x}).$$

We next prove a result which is fundamental in all our subsequent discussions of specializations of Φ.

LEMMA 2. *If $M(\mathfrak{x})$ is any non-zero polynomial element of Φ such that $M^G = M(T_G\mathfrak{x}) = m_G M(\mathfrak{x})$ for every G in $\mathfrak{G}$, where m_G is an element of P, then there is some $\mathfrak{p}$, not satisfying any equation of the form $T_G\mathfrak{p} = \mathfrak{p}$ except when G is the unit element E of $\mathfrak{G}$, for which $M(R\mathfrak{p})$ is different from zero.*

Proof. Let us suppose that $M(R\mathfrak{p}) = 0$ for every $\mathfrak{p}$ which does not satisfy an equation of the form stated. Let us as before assume that $F = P(\theta)$, where θ is an element of F, and let us write θ_G for θ^G. Then if

$$H(\mathfrak{p}) = \Sigma_G(\theta^j M(R\mathfrak{p}))^G,$$

where G runs through the elements of $\mathfrak{G}$, and j is a positive integer to be suitably chosen, then

$$H(\mathfrak{p}) = \Sigma_G \theta_G{}^j M^G(R\mathfrak{p}) = \Sigma_G \theta_G{}^j M(T_G R\mathfrak{p}) = (\Sigma_G \theta_G{}^j m_G) M(R\mathfrak{p}).$$

The coefficient of $M(R\mathfrak{p})$ on the right side is different from zero if j is suitably chosen. The expression $H(\mathfrak{p})$ has elements in P which vanish for every $\mathfrak{p}$ which does not satisfy any equation of the form stated. If we observe that $\Pi'_G((T_G - I)\mathfrak{p})H(\mathfrak{p})$, where G runs through all elements except the unit element of $\mathfrak{G}$, is a matrix which vanishes for every $\mathfrak{p}$, we easily conclude that $H(\mathfrak{p})$ vanishes for every $\mathfrak{p}$. Therefore the coefficients of $p_1, p_2, \cdots, p_q$ in $H(\mathfrak{p})$ must vanish, and consequently the coefficients of $p_1, p_2, \cdots, p_q$ in $M(R\mathfrak{p})$ must vanish. It follows that we may replace $\mathfrak{p}$ in $M(R\mathfrak{p})$ by any matrix of order q with zeros in the 2-nd, 3-rd, $\cdots$, q-th columns, and the resulting expression must vanish. We therefore replace $\mathfrak{p}$ by $R^{-1}\mathfrak{x}$, which exists since R is non-singular (by Lemma 1), and we conclude that $M(\mathfrak{x}) = 0$ for every $\mathfrak{x}$. But this is a contradiction. Therefore the lemma is true.

From our preliminary study of conjugate double systems we obtain

COROLLARY 1. *If $\zeta_{A,B}$ is a conjugate double system in Φ, then there exists a specialization of Φ into F such that the system $\zeta_{A,B}$ in Φ becomes a well defined non-zero conjugate double system $z_{A,B}$ in F.*

COROLLARY 2. *If $f(\mathfrak{x})$ is a non-zero element of Π, then there exists a specialization of Φ into F such that $f(\mathfrak{x})$ becomes a well defined non-zero element of P.*

For we need only take M in the lemma [11] to be the product of the numerator and denominator of $f(\mathfrak{x})$. The hypotheses of the lemma are easily seen to be satisfied, by an argument analogous to that used in the discussion of conjugate double systems.

From the first corollary we obtain

COROLLARY 3. *Given any non-zero element σ of Σ, there is a specialization of Φ into F such that σ becomes a well defined non-zero element s of S.*

Finally, we also have

COROLLARY 4. *Given any finite number of conjugate doubte systems in Φ, any finite number of non-zero elements of Π, and any finite number of non-zero elements of Σ, there exists a specialization of Φ into F such that*

[11] Or we may regard Corollary 2 as a consequence of Corollary 1.

these become well defined conjugate double systems in F, non-zero elements of P, and non-zero elements of S respectively.

For we can take M in the lemma to be the product of all the non-zero numerators and denominators of all the elements of Φ which enter into the definition of these expressions.

We can now consider some of the properties of the algebra $\mathfrak{S}$.

THEOREM 4. *If ϵ is the exponent of the factor set c in the field Φ, and e is the exponent in the field F, then e is a divisor of ϵ.*

Proof. The number ϵ is the smallest positive integer such that there exists a conjugate double system $\zeta_{A,B}$ in Φ, such that

$$c^{\epsilon}{}_{A,B,C} = \frac{\zeta_{A,B}\zeta_{B,C}}{\zeta_{A,C}}.$$

Let us now, using Corollary 1 of Lemma 2, choose a specialization of Φ into F by which the system $\zeta_{A,B}$ becomes a conjugate double system $z_{A,B}$ in F. Then

$$c^{\epsilon}{}_{A,B,C} = \frac{z_{A,B}z_{B,C}}{z_{A,C}}.$$

Therefore e divides ϵ, since e is the minimum exponent with this property.

THEOREM 5. *The index m of S is a divisor of the index μ of $\mathfrak{S}$.*

Proof. Suppose that D,Δ are the Wedderburn factors of S,$\mathfrak{S}$ respectively, so that

$$S \cong D \times [P]_t, \text{ and } \mathfrak{S} \cong \Delta \times [\Pi]_\tau,$$

where $mt = r = \mu\tau$. We may write $\Delta = \delta\mathfrak{S}\delta$, where δ is a primitive idempotent element of $\mathfrak{S}$. Let us now choose a specialization of Φ to F such that δ becomes a well defined element d of S. Then Δ becomes an algebra $D_1 = dSd$ over P. Also, $\mathfrak{S}$ becomes S, and $[\Pi]_\tau$ becomes $[P]_\tau$. Therefore

$$S \cong D_1 \times [P]_\tau.$$

Since $[P]_\tau$ and S are central and simple, D_1 also is central and simple.[12] If now the Wedderburn factor of D_1 is D_2, then $D_1 \cong D_2 \times [P]_u$, where u is some positive integer, and consequently

$$S \cong D_1 \times [P]_\tau \cong D_2 \times [P]_u \times [P]_\tau \cong D_2 \times [P]_{u\tau}.$$

---

[12] See A. A. Albert, "On direct products, cyclic division algebras and pure Riemann matrices," *Transactions of the American Mathematical Society*, vol. 33 (1931), p. 221.

It follows then, from Wedderburn's theorem, that $D_2 \cong D$ and $u\tau = t$. From the relation $mt = \mu\tau$, we find that $\mu = um$. Therefore m is a divisor of μ.

THEOREM 6. *If Σ has a normal splitting field Ω of degree s over Π, then S has a normal splitting field V of which the degree over P is a divisor of s. The Galois group of V over P is (isomorphic to) a sub-group of the Galois group of Ω over Π.*

Proof. Let $\mathfrak{W}$ be the Galois group of Ω over Π, and let $\Omega = \Pi(\omega)$, where ω is a root of the separable and irreducible equation

$$\phi(t) = t^s + \alpha_1 t^{s-1} + \cdots + \alpha_s = 0.$$

The coefficients $\alpha_1, \alpha_2, \cdots, \alpha_s$ are elements of Π. Let δ be the discriminant of $\phi(t)$, and let $\omega_E = \omega, \omega_A, \omega_B, \cdots$ be the conjugates of ω, where E is the unit element, and $A, B, \cdots$ are the other elements of the Galois group $\mathfrak{W}$. Then there are relations of the form

$$\omega_A = \gamma_A(\omega) \text{ and } \gamma_A(\gamma_B(\omega)) = \gamma_{AB}(\omega)$$

for every A,B of $\mathfrak{W}$, where $\gamma_A(t)$ is a polynomial of degree less than s, with coefficients in Π.

Let $\lambda_1(t), \lambda_2(t), \cdots, \lambda_s(t)$ be the elementary symmetric functions of degree $1, 2, \cdots, s$ respectively in the s functions of the form $\gamma_A(t)$ for A in $\mathfrak{W}$. Then

$$\lambda_i(\omega) = (-1)^i \alpha_i \qquad (i = 1, 2, \cdots, s).$$

Since $\phi(t)$ is irreducible, there exist polynomials $\nu_{A,B}(t)$ for every pair of elements A,B of $\mathfrak{W}$, and polynomials $\kappa_i(t)$ for $i = 1, 2, \cdots, s$, such that

$$\gamma_A(\gamma_B(t)) = \gamma_{AB}(t) + \nu_{A,B}(t)\phi(t)$$

and

$$\lambda_i(t) = (-1)^i \alpha_i + \kappa_i(t)\phi(t).$$

Let us now choose a specialization of Φ into F such that δ and all the coefficients of all the polynomials ϕ, γ_A, $\nu_{A,B}$ and κ_i become non-zero elements of P. Then if $\phi(t)$ becomes $f(t) = t^s + a_1 t^{s-1} + \cdots + a_s$, and $\gamma_A(t)$, $\nu_{A,B}(t)$, $\lambda_i(t)$, $\kappa_i(t)$ become $g_A(t)$, $n_{A,B}(t)$, $l_i(t)$, $k_i(t)$ respectively, we have

$$g_A(g_B(t)) = g_{AB}(t) + n_{A,B}(t)f(t) \qquad (A,B \text{ in } \mathfrak{W}),$$

and

$$l_i(t) = (-1)^i a_i + k_i(t)f(t) \qquad (i = 1, 2, \cdots, s).$$

Let w be a root of the equation $f(t) = 0$, and let $V = P(w)$. Further, let

$\psi(t)$ be an irreducible factor of $f(t)$ such that $\psi(w) = 0$. Then the above relations show that each root of the equation $\psi(t) = 0$ can be written as $g_A(w)$, where A is an element of a subset $\mathfrak{V}$ of $\mathfrak{W}$. Since δ becomes a non-zero element of P, $f(t)$ and therefore also $\psi(t)$ must be separable. We have thus seen that $\psi(t)$ is normal, separable, and irreducible. Automorphisms of V over P may be defined by the transformation definition $w^A = g_A(w)$, for A in $\mathfrak{V}$, since $g_A(g_B(w)) = g_{AB}(w)$. Evidently $\mathfrak{V}$ is a subgroup of $\mathfrak{W}$, and V is normal over P with Galois group $\mathfrak{V}$. Now Ω may be considered to be the field of all polynomials, modulo $\phi(t)$, with coefficients in Π. Suppose $\beta(t)$ is any polynomial with coefficients in Π, which specializes to a polynomial $b(t)$ with coefficients in P. Then division by $\phi(t)$ gives a relation

$$\beta(t) = \phi(t)\pi(t) + \rho(t),$$

where $\pi(t)$ and $\rho(t)$ are polynomials with coefficients which are integral rational functions of the coefficients of $\beta(t)$ and $\phi(t)$, and where the degree of $\rho(t)$ is less than s. Therefore $\pi(t)$ and $\rho(t)$ must specialize to polynomials $p(t)$ and $r(t)$ respectively, and we have

$$b(t) = f(t)p(t) + r(t).$$

Consequently Ω specializes to a commutative algebra W consisting of all polynomials, modulo $f(t)$, with coefficients in P. Since $f(t)$ is in general reducible, W is in general not a field. The field Ω being a splitting field for Σ, there must exist in Ω an absolutely irreducible representation Λ of degree r. In this representation the element κ of Π may be assumed to be mapped on κI. For, if Q is the image of κ, and if k is a characteristic root of the matrix Q, then $\Lambda(Q - kI) = (Q - kI)\Lambda$, and consequently, by Schur's lemma, $Q - kI = 0$, that is, $Q = kI$. If $\kappa \neq k$, the correspondence $k \rightarrow \kappa$ defines an automorphism of P. By applying to Λ the inverse automorphism we obtain an absolutely irreducible representation in which the image of κ is κI. Let $\sigma_1, \sigma_2, \cdots, \sigma_{r^2}$ form a basis of Σ over Π, let $\Lambda_1, \Lambda_2, \cdots, \Lambda_{r^2}$ be their representatives in Λ, and let $\Lambda_v = (\lambda_{i,j}^{(v)})$ for $v = 1, 2, \cdots, r$. Then $\lambda_{i,j}^{(v)}$, being an element of Ω, may be considered to be a polynomial of degree less than s, with coefficients in Π. Let us choose a specialization of Φ into F in such a way that the requirements of the first part of our argument are satisfied, and also in such a way that all the coefficients in the polynomials $\lambda_{i,j}^{(v)}$ become non-zero elements of P. Then $\Lambda_1, \Lambda_2, \cdots, \Lambda_{r^2}$ become matrices with coefficients in W, which define a representation of S. But since we may obtain V from W by identifying elements of W of which the difference is divisible by $\psi(t)$, the representation of S in

W determines a representation of S in V. Since this representation in V has degree r, it must be absolutely irreducible, and therefore V must be a splitting field of S.

Since a subgroup of a cyclic group is cyclic, we obtain the

COROLLARY. *If Σ has a cyclic splitting field of degree s over Π, then S has a cyclic splitting field of which the degree over S is a divisor of s.*

5. The results of the previous sections have been derived without making any special assumptions about the splitting field F and the factor set c of S. In order to obtain more precise results we now assume that the splitting field F is regular;[13] that is, we assume that F has been constructed in such a way that the elements of the factor set c are all roots of unity.

Suppose that $F_0 = P(\eta)$ is any separable splitting field of S of degree v over P. Let F_1 be the normal field obtained by adjoining all the conjugates of η to P, and let $\mathfrak{G}_1$ be the Galois group of F_1 over P. In F_1 we can construct a factor set c_1 with exponent e_1. If now we adjoin to F_1 the n-th roots of suitable elements of F_1, where n is a multiple of e_1, we obtain a regular field F. In F the factor set c_1 is associated with a factor set c of which all the elements are n-th roots of unity. (If $\mathfrak{G}$ is the Galois group of F over P, then the definition of c_1 is first extended to the group $\mathfrak{G}$ by the natural homomorphism of $\mathfrak{G}$ on $\mathfrak{G}_1$.) Let us now denote by $\mathfrak{N}$ the group of n-th roots of unity in F. Then the rules

$$Z_A Z_B = Z_{ABC^{-1}{}_{AB,B,E}}$$

$$\lambda Z_A = Z_A \lambda^A,$$

where E is the unit element, and A,B are any elements of $\mathfrak{G}$, and λ is in $\mathfrak{N}$, define a multiplicative group $\mathfrak{H}$ consisting of all products of the form λZ_A. The associativity of multiplication is secured by the property (3) of factor sets which was mentioned in Section **1**. The group $\mathfrak{N}$ is an invariant subgroup of $\mathfrak{H}$, and $\mathfrak{H}/\mathfrak{N} \cong \mathfrak{G}$.

The assumption made at the beginning of Section **2**, namely, that P contains the elements of the factor set c, is now satisfied if we make the weaker and at the same time more simple assumption that P contains the n-th roots of unity.

---

[13] For the details of this theory, of which only an outline is given here, see R. Brauer, "Über die Konstruktion der Schiefkörper, die von endlichem Rang in bezug auf ein gegebenes Zentrum sind," *Journal für die Reine und Angewandte Mathematik*, vol. 168 (1932), pp. 59-60.

In terms of the group $\mathfrak{H}$, defined by the regular splitting field F, we now state [14] a theorem due to R. Brauer by which the exponent ϵ is determined explicitly.

THEOREM 7. *If the factor set c lies in P, and if $\mathfrak{H}$ is the group associated with c, then the exponent with respect to Φ and Π of the factor set c is the smallest positive number ϵ for which there is a subgroup $\mathfrak{K}$ of $\mathfrak{H}$ with the following properties*: 1) *$\mathfrak{K}$ contains the commutator subgroup of $\mathfrak{H}$.* 2) *$\mathfrak{K}$ has an intersection of order ϵ with the group $\mathfrak{N}$ of n-th roots of unity.* 3) *There is a number t, such that $\mathfrak{K}$ contains the t-th powers of all group elements of $\mathfrak{H}$ and at the same time P contains the t-th roots of unity.*

We next employ the regular field F and the group $\mathfrak{H}$ associated with c to obtain a more precise result about the index μ.

Let $\mathfrak{S}$ be the subgroup of $\mathfrak{G}$ which belongs to the subfield F_0 of F. Then [15] $c_{A,B,C} = 1$ if A,B,C are in $\mathfrak{S}$, and consequently the elements Z_S of $\mathfrak{H}$ with S in $\mathfrak{S}$ form a subgroup $\mathfrak{R}$ of $\mathfrak{H}$ which has only the unit element of $\mathfrak{H}$ in common with $\mathfrak{N}$. The degree of F_0 over P being v, the order of $\mathfrak{S}$, and therefore also of $\mathfrak{R}$ must be r/v. According to an established property [16] of the group $\mathfrak{H}$, the index of any algebra possessing the factor set with which $\mathfrak{H}$ is associated is a divisor of $r/(r/v)$, i. e. of v. Since S has a splitting field of order equal to its index m, we may take $v = m$, and since the algebra Σ possesses the same factor set as does A, the above argument shows that the index μ of Σ is a divisor of m. Combining this result with Theorem 5, we see that we have thus proved

THEOREM 8. *If the regular splitting field F is constructed by extending a splitting field of S of degree m over P in the way described, then $\mu = m$.*

An algebra which can be defined, as we have defined S over P and Σ over Π, by means of a normal field and a factor set contained in the field, is called a crossed product of the normal field and its Galois group. Let us suppose that S has a normal splitting field V of degree m over P. Then we may define a factor set of S in V, and with this factor set and the field V we may define a crossed product algebra B. Since the order of B is m^2, B must be isomorphic to D, the Wedderburn factor of both S and B. That is, if A has a normal splitting field of degree m over P, then D is a crossed

[14] Without proof. See R. Brauer, "Über die Kleinsche Theorie der Algebraischen Gleichungen," *Mathematische Annalen*, vol. 110 (1934), p. 500.

[15] See R. Brauer, *loc. cit.*[13], p. 60.

[16] See R. Brauer, *loc. cit.*[13], p. 58.

product. Similarly, if Σ has a normal splitting field of degree μ over Π, then Δ is a crossed product. From Theorems 6 and 8 we therefore obtain

THEOREM 9. *If the hypotheses of Theorem* 8 *are true, and if* Δ *is a crossed product, then* D *is a crossed product.*

Crossed products with a cyclic field, that is, cyclic algebras, are of particular interest. From the corollary to Theorem 6 we derive the

COROLLARY. *If the hypotheses of Theorem* 8 *are true, and if* Δ *is a cyclic algebra, then* D *is a cyclic algebra.*

From Theorem 9 we obtain directly [17]

THEOREM 10. *If there exist central division algebras which are not crossed products, then there exist generic algebras which are not crossed products.*

THEOREM 11. *If the hypotheses of Theorem* 8 *are true, and if* Δ *is a direct product, then* D *is a direct product.*

Proof. We have seen, in the proof of Theorem 5, that Δ specializes to an algebra $D_1 = D \times [P]_u$, where $\mu = um$. By Theorem 8, $u = 1$ and consequently $D_1 = D$. Suppose that $\Delta = \Delta_2 \times \Delta_3$, where Δ_2, Δ_3 are subalgebras of Δ. Let $\eta_1, \eta_2, \cdots, \eta_h$ be a basis for Δ_2 over Π, and let $\zeta_1, \zeta_2, \cdots, \zeta_k$ be a basis for Δ_3 over Π. Each η commutes with each ζ. We now choose a specialization such that the elements $\eta_1, \eta_2, \cdots, \eta_h$ and the elements $\zeta_1, \zeta_2, \cdots, \zeta_k$, which are in Δ, and therefore also in Σ, become well defined systems y and z in S. Evidently each y commutes with each z. Therefore Δ specializes to $D = D_2 \times D_3$, where D_2 and D_3 are central simple [18] subalgebras of D.

UNIVERSITY OF SASKATCHEWAN,
SASKATOON, CANADA.

[17] It is not yet known whether there exist central division algebras which are not crossed products. This theorem might afford a means of discussing this problem.

[18] See A. A. Albert, *loc. cit.*

SIMPLE LINKS AND FIXED SETS UNDER CONTINUOUS MAPPINGS.*

By J. L. KELLEY.

In generalizing a result of W. L. Ayres [1] I have shown [2] that a homeomorphism of any compact continuum into itself leaves invariant a certain sub-continuum which is without cut points. From this conclusion it follows that under a homeomorphism either there is a fixed point or some 0-th order cyclic element is carried into a subset of itself (according to G. T. Whyburn, a 0-th order cyclic element is a continuum maximal with respect to the property of containing no cut points of itself [3]). In attempting to obtain related results for any continuous transformation R. L. Moore's decomposition into *simple links* is found to be more suitable.[4] In what follows new results on simple links are proven, and in logically developing the theory some of Moore's results are repeated.

The last section is devoted to transformation theory. It is shown that for any transformation T of a compact continuum M into itself, there exists a continuum π, a subset of a simple link of M, such that $T(\pi) \supset \pi$. Corollaries to this theorem include the theorem of Ayres previously men-

* Received October 2, 1946; Presented to the Society December, 1938. This manuscript was offered to *Fundamenta Mathematicae* in June, 1939. An abstract appeared in *Proceedings of the National Academy of Sciences*, vol. 26 (1940), pp. 192-194. Two changes in the nomenclature of the abstract have been made, namely, instead of *F*-set, simple link; instead of *B*-set, central set.

Since this paper was written a unified treatment of cyclic element theory has appeared in G. T. Whyburn's *Analytic Topology*, American Mathematical Society Colloquium Publications. Certain of the results of the first sections of the paper appear in this book.

The author has benefited considerably from discussions of these topics with Professor Whyburn.

[1] See *Fundamenta Mathematicae*, vol. 16 (1930), pp. 333-336.

[2] See *Duke Mathematical Journal*, vol. 5 (1939), pp. 535-537.

[3] See *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 159-165.

[4] See *Foundations of point set theory*, American Mathematical Colloquium Publications, p. 72. The definition of simple link given here differs from Moore's only in that all cut points are simple links. This change is made to conform to cyclic element theory for Peano curves.

The first half of 2.1 and Theorems 2.4 and 2.5 of this paper are contained in Moore's results. The methods of proof differ considerably from Moore's.

tioned and the Scherrer theorem [6] that any dendrite has the fixed point property.

1. Preliminaries. Throughout, M will denote a compact metric continuum.

Definition.[7] A continuum $A \subset M$ is a *nodal set* if $A \cdot \overline{M - A}$ consists of at most a single point. A continuum B is a *central set* if $\overline{M - B}$ consists of a finite number of components, each cutting B in a single point. Note that M itself is a central set.

We shall need the following properties of nodal sets:

1.1. a) *If A is a nodal set so is $\overline{M - A}$.*

b) *If A is a nodal set and C is connected, then $A \cdot C$ is connected or vacuous.*

c) *The union of two intersecting nodal sets is a nodal set.*

Proof. Both a) and b) are simple consequences of the definition. Suppose that A_1 and A_2 are intersecting nodal sets and that two distinct points p, $q \in (A_1 \cup A_2) \cdot \overline{(M - (A_1 \cup A_2))}$. Since $A_1 \cdot \overline{(M - A_1)}$ and $A_2 \cdot \overline{(M - A_2)}$ each consist of a single point we must have $\{p\} \cup \{q\} = (A_1 \cup A_2) \cdot \overline{(M - A_1)} \cdot \overline{(M - A_2)}$. From b), however, it is clear that the intersection of a connected set and two nodal sets is connected, and we have a contradiction.

1.2. *In order that the closed set B be a central set it is necessary and sufficient that $\overline{M - B}$ be the union of a finite number of nodal sets, or, equivalently, that B be the intersection of a finite number of nodal sets.*

Proof. The necessity is clear. By 1.1 c) any finite union of nodal sets can be reduced to a finite union of non-intersecting nodal sets and the sufficiency follows. From 1.2 we have immediately:

1.3. *The intersection of a finite number of central sets is a central set.*

If each of a system of sets in M has the property that its intersection with a continuum is a continuum or vacuous, then the intersection of all the sets of the system also has this property. Hence,

---

[6] See *Mathematische Zeitschrift*, vol. 24 (1926), p. 129.

[7] The definition of nodal set is due to G. T. Whyburn, *Duke Mathematical Journal*, vol. 4 (1938), p. 2.

1.4. *If C is a continuum and A is the intersection of central sets then $A \cdot C$ is a continuum or vacuous.*

Definitions.[8] A point $p \in M$ is *conjugate* to $q \in M$ if no point separates p and q in M. If $p \in M$, M_p is defined to be the set of all points conjugate to p. A point $p \in M$ is an *end point* of M if there exists an arbitrarily small neighborhood of p having as its boundary a single point. A set is said to be a simple link provided

a) it consists of an end point or a cut point of M, or

b) it is a *non-degenerate* M_p, where p is a non-cut point.

The latter sets are called *true* simple links.

1.5. *Any set M_p is the monotone intersection of central sets.*

Proof. The set M_p is clearly the product of all nodal sets containing p in their interior. This product can be written as the product of a countable number of nodal sets, and every partial product is a central set. The result follows.

2. Simple links. From the preceding section we have at once the following:

2.1. THEOREM. *A simple link is a continuum. The intersection of a simple link and a continuum is a continuum or vacuous.*

2.2. THEOREM. *If p_1 and p_2 do not belong to M_q and cannot be joined by a continuum in $M - \{q\}$ then M_q separates M between p_1 and p_2.*

Proof. Write $M_q = \cdot B_n$, where $\{B_n\}$ is a monotone sequence of central sets each containing q in its interior. For each $p \in M - M_q$ let N_p be the union of the components about p of $\overline{M - B_n}$. For each p, $N_p + M_q$ is closed, for if N_p intersects a component of $\overline{M - B_n}$ it contains this component. If p_1 and p_2 are the points of the hypothesis of the theorem, the sets N_{p_1} and N_{p_2} are disjoint, for p_1 can be joined to any point of N_{p_1} by a continuum in $M - \{q\}$, and similarly for N_{p_2}. Finally, almost all the sets N_p lie within any B_n, since $\overline{M - B_n}$ has a finite number of components. Hence, $M - M_q = (N_{p_1} - M_q) + (M - N_{p_1} - M_q)$ is a separation between p_1 and p_2.

[8] See K. Kuratowski and G. T. Whyburn, *Fundamenta Mathematicae*, vol. 16 (1930), pp. 305-331. R. L. Moore defines the sets M_p (ref. 4).

2.3. THEOREM. *If p is neither a cut point nor an end point then there exists a point $q \neq p$, q conjugate to p, i.e., M is the union of its simple links.*

Proof. Suppose that p has no conjugate point and that it is not a cut point. Write $\{p\} = M_p = \cdot B_n$, where B_n is a monotone decreasing sequence of central sets. For $\epsilon > 0$ choose m so that dia $B_m < \epsilon$. Since p is not a cut point and $M_p = \{p\}$ we may, by the previous theorem, join the components of $\overline{M - B_m}$ to each other by continua in $M - \{p\}$. Let A be a continuum $p \bar{\epsilon} A$, $A \supset \overline{M - B_m}$. For m_1 sufficiently large, $B_{m_1} \cdot A = 0$ and the complement of the component about A of $\overline{M - B_{m_1}}$ is a neighborhood of p of diameter $< \epsilon$ with a single boundary point. It follows that p is an end point.

2.4. THEOREM. *Each non-cut point of M belongs to one and only one simple link.*

Proof. We need only consider $p \in M_q \cdot M_r$, where M_q and M_r are true simple links. If $M_q - M_r \neq 0$ there exists a nodal set A containing M_q but not M_r. Then $\overline{M - A}$ is a nodal set containing M_r and $A \cdot \overline{M - A}$ must be p, and p is then a cut point.

2.5. THEOREM. *In order that two points p and q belong to the same simple link it is necessary and sufficient that p and q be conjugate.*

Proof. If p and q belong to the same simple link they are surely conjugate. Suppose p is conjugate to q, $p \neq q$. Let $A = M_p \cdot M_q$. Now A is a continuum, and since M_p can be separated in M by no point save possibly p, and similarly for M_q, it follows that A cannot be separated in M by any point. No cut point in A can then be of order 2, and hence [9] we can choose a non-cut point $r \in A$. Clearly, p, $q \in M_r$ and the theorem is proved.

2.6. THEOREM. *If a point p does not belong to a true simple link it is a regular point in the sense of Menger-Urysohn.*

Proof. Write $\{p\} = M_p = \cdot B_n$, where $\{B_n\}$ is a monotone sequence of central sets, each containing p in its interior. Then $\{B_n - \overline{M - B_n}\}$ is the required sequence of neighborhoods.

Since no cut point belonging to a true simple link can be of order 2 we have

---

[9] The order of a point p is the least integer n such that there exists an arbitrarily small neighborhood of p with at most n boundary points. See K. Menger, *Kurventheorie*, p. 96. All but a countable number of cut points of a continuum are of order 2. See G. T. Whyburn, *Transactions of the American Mathematical Society*, vol. 30 (1928), pp. 597-609.

2.7. THEOREM. *There are in M only a countable number of cut points belonging to true simple links.*

2.8. THEOREM. *In order that a non-degenerate closed subset of M be a true simple link it is necessary and sufficient that it be separated in M by no point of M, and that it be saturated relative to this property.*

Proof. The necessity is clear. Suppose that A is saturated relative to the property. If $p, q \in A$, p is conjugate to q and hence by 2.5, $p.q \in M_r$, a true simple link. Since any point not in M_r fails to be conjugate to one of p or q, $A \subset M_r$. Since M_r is not separated in M by any point of M, $M_r = A$.

2.9. THEOREM. *In order that a closed set $A \subset M$ be a continuum it is necessary and sufficient that $A \cdot F$ be a continuum for every simple link F and that A contain every point separating a pair of its points in M.*

Proof. The necessity is clear. Let A be a closed set satisfying the two given conditions and suppose $A = A_1 \cup A_2$ is a separation. Let A^* be a continuum irreducible with respect to the property of cutting both A_1 and A_2. Then A^* must be irreducible about two points $p \in A_1 \cdot A^*$, $q \in A_2 \cdot A^*$. If p is conjugate to q and F is the simple link about p and q then $A \cdot F$ is connected and $A_1 \cup A_2$ cannot be a separation between p and q. If $r \in M$ separates p and q in M then $r \in A^*$ and A^* cannot be irreducible between A_1 and A_2. We have a contradiction.

PROBLEM. It is of interest to determine the relation between the simple link decomposition and Whyburn's E_0-set decomposition. Each simple link F is itself a continuum and we may decompose F into simple links $\{F^1\}$ in precisely the same way we decompose M. Proceeding inductively we define F^α for every ordinal α.

1) if α has a predecessor $\alpha - 1$ the sets F^α are the simple links of $F^{\alpha-1}$ sets.

2) if α is a limit number the sets F^α are all possible intersections $\cdot F^\beta$, where this product contains a set from every class $\{F^\beta\}$, for $\beta < \alpha$.

Now a set is both an F^α and an $F^{\alpha+1}$ if and only if it contains no cut point of itself. Any E_0-set is a subset of some F^α for every α. Hence, any $E_0 = F^\beta$ for some β in the first or second number class. (We can then deduce that the product of an E_0-set and a continuum is a continuum or vacuous.) The problem is

For a given M is there a number $\beta = \beta(M)$ in the first or second number class such that the class of sets F^β is identical with the class of E_0-sets?

3. Generalizations of A-sets; Cyclically reducible and extensible properties. In a manner analogous to a characterization of an A-set for (see ref. 8) Peano continua we define

Definition. A subcontinuum of M is a *J-set* if it is the union of simple links.

We remark that not all J-sets are intersections of central sets.

We have immediately from 2.9 and 2.1 the

3.1. THEOREM. *The intersection of a continuum with a J-set is a continuum or vacuous.*

The following theorems are now obvious.

3.2. THEOREM. *The intersection of any number of J-sets is a J-set, or vacuous.*

3.3. THEOREM. *A J-set contains all the irreducible continua about any two of its points.*

In generalizing the notion of cyclic chain it is most natural to define the cyclic chain $C(a, b)$ as the intersection of all J-sets containing a and b. However, much of the usefulness of the cyclic chain is lost. If $E(a, b)$ is the set of all points separating a and b in M, then every true simple link in $C(a, b)$ has at least one point in common with the closure of $E(a, b) \cup \{a\} \cup \{b\}$; there may be only one common point. Every true simple link in $C(a, b)$ separates a and b in M; there may be, however, cut points in $C(a, b)$ belonging to no true simple link of $C(a, b)$ which fail to separate a and b in M. A characterization of $C(a, b)$ in terms of $E(a, b)$ is too complicated to be useful.

We define in the usual way (see ref. 8)

Definition. A property is *simple link extensible* if when every simple link of M has the property so has M. A property is *simple link reducible* if when M has the property so also has every simple link of M.

3.4. THEOREM. *Unicoherence is simple link extensible (but not necessarily simple link reducible).*

Proof. If $M = M_1 \cup M_2$ where M_1 and M_2 are continua and if every simple link is unicoherent, then $M_1 \cdot M_2$ satisfies the hypothesis of 2.9 and

is hence connected. On the other hand the plane set described in polar coordinates by $\rho = 1$ for $0 \leqq \theta \leqq 2\pi$ and $\rho = 1 + 1/\theta$ for $\theta \geqq 1$ is a unicoherent continuum containing a non-unicoherent simple link.

The relation of the space to its simple links is naturally more complicated in general than is the case for Peano continua. By means of examples it can be shown that

3.5. a) *A simple link (E_0-set) is not necessarily a retract of the space.*[10]

b) *The fixed point property for homeomorphisms (i.e., the property that under any homeomorphism of the space into itself at least one point remains fixed) is not simple link reducible.*

However, the following will be a direct consequence of 4.8:

3.6. THEOREM. *The fixed point property for homeomorphisms is simple link extensible.*

4. Fixed sets under continuous mappings. Throughout this section T will denote a given continuous map of M into M. The following has been shown by the author in a previous paper (see ref. 2).

4.1. *If $T: M \to M$ is continuous then $\pi = \cdot \, T^i(M)$, $i = 1, \cdots$, is a non-vacuous continuum fixed under T (i.e., $T(\pi) = \pi$).*

4.2. *The property of a closed set A of being contained in its transform $T(A)$ is inducible.*

Proof. Suppose that $A = \cdot \, A_i$, where $T(A_i) \supset A_i$ and $A_i \supset A_{i+1}$. If $p \in A$ then since $p \in A_i$ for all i and $T(A_i) \supset A_i$ it must be true that $T^{-1}(p) \cdot A_i \neq 0$ for all i. Hence, $T^{-1}(p) \cdot A \neq 0$ and $p \in T(A)$.

Definitions. A continuum $A \subset M$ *has property P* if

a) it is a central set, and

b) if B is a component of $\overline{M - A}$, with $B \cdot (\overline{M - A}) = \{p\}$, then $T(p) \bar{\in} B - \{p\}$.

Notice that M itself has P, since any component of $\overline{M - M}$ surely satisfies b).

A continuum in M *has property P^** if it is the intersection of sets having property P. The property P^* is clearly inducible.

4.3. a) *If A_1 and A_2 have P and $A_1 \cdot A_2 \neq 0$ then $A_1 \cdot A_2$ has P.*

---

[10] A set $A \subset B$ is a retract of B if there is a continuous map of B onto A which is the identity on A. See K. Borsuk, *Fundamenta Mathematicae*, vol. 17 (1931), pp. 152-170.

b) *If A is irreducible with respect to the property P^* then A is a simple link. If A is a true simple link it may, by* a), *be written as the intersection of a monotone family of sets having P.*

Proof. We first prove a). Suppose that B_1 and B_2 are nodal sets with $\{p_i\} = B_i \cdot \overline{M - B_i}$, $T(p_i) \bar{\varepsilon} B_i - \{p_i\}$, $(i = 1, 2)$ and $B_1 \cdot B_2 \neq 0$. Then by 1.2c), $B_1 + B_2$ is nodal and we can suppose that $\{p_1\} = \overline{(B_1 + B_2) - (B_1 + B_2)}$. If $p_2 \neq p_1$ then $B_2 \subset B_1$ and $T(p_1) \bar{\varepsilon} B_1 + B_2 - \{p_1\}$. If $p_1 = p_2$, clearly we again have $T(p_1) \bar{\varepsilon} B_1 + B_2 - \{p_1\}$. Making use of this addition of nodal sets and of 1.4 the truth of a) is clear.

We now prove b). Suppose that $A = \cdot A_i$ is a set irreducible with respect to P^*, and that $\{A_i\}$ is a sequence of sets with P. If B is any nodal set, then either B or $\overline{M - B}$ has P. If B is nodal and has P, and $B \cdot A \neq 0$ then $B \cdot A = \cdot (B \cdot A_i)$ has P^*. Hence, A is separated in M by no point of M and is therefore a subset of a simple link. But A is the intersection of central sets and must therefore be a simple link.

Definition. For $A \subset M$ we define $A_1 = A \cdot T(A)$ and inductively $A_n = A \cdot T(A_{n-1})$. Let $D(A) = \cdot A_n$, $n = 1, \cdots$. If $A_0 \subset A$ clearly $D(A_0) \subset D(A)$.

4.4. *For any A having P, $D = D(A)$ is a non-void continuum and $T(D) \supset D$.*

Proof. Write $\overline{M - A} = + B_i$, $i = 1, \cdots n$, where B_i are pairwise disjoint continua with $B_i \cdot A = \{p_i\}$, $T(p_i) \bar{\varepsilon} B_i - \{p_i\}$. For $p \varepsilon A$ we define a map $\phi(p)$ as follows

$$\text{if } T(p) \varepsilon A \text{ then } \phi(p) = T(p)$$
$$\text{if } T(p) \varepsilon B_i \text{ then } \phi(p) = p_i.$$

The continuity of ϕ is readily verified. Now $\phi(A) = A_1$, and it is a simple induction proof to show that $\phi^n(A) = A_n$. By 4.1, $D = \cdot \phi^n(A)$ is a continuum fixed under ϕ. We must show that $T(D) \supset D$. If D consists of a single point, in view of the definition of ϕ, this point is surely fixed under T. If D is non-degenerate $T(D) \supset D - + \{p_i\}$, $i = 1, \cdots, n$, and hence $T(D) \supset D$.

4.5. THEOREM. *If $T: M \to M$ is a continuous transformation then there exists a continuum π, a subset of a simple link of M, such that $T(\pi) \supset \pi$.*

Proof. Let $A = \cdot A_i$ be irreducible with respect to P^* where $\{A_i\}$ is a monotone decreasing sequence of sets with P. Let $\pi = \cdot D(A_i)$. The sequence $\{D(A_i)\}$ is monotone decreasing and $T[D(A_i)] \supset D(A_i)$. The theorem then follows from 4.2 and 4.3 b).

If π is degenerate we have a fixed point. Hence

4.6. THEOREM. *If $T: M \to M$ is continuous, there is either a fixed point in M or a simple link F such that $F \cap T(F)$ is a non-degenerate continuum.*

4.7. THEOREM. *If $T: M \to M$ is continuous there exists a compact subset A of a simple link of M such that $T(A) = A$.*

Proof. Let π be the set of 4.5. Define $A_1 = \pi \cap T^{-1}(\pi)$, and then inductively for each n, $A_n = \pi \cap T^{-1}(A_{n-1})$. If $A_{n-1} \subset A_{n-2}$ then $A_n = \pi \cap T^{-1}(A_{n-1}) \subset \pi \cap T^{-1}(A_{n-2}) = A_{n-1}$ and hence $\{A_n\}$ is a monotone sequence of sets. For any $p \in \pi$, $T^{-1}(p) \cap \pi \neq 0$ and hence no $A_n = 0$.

Let $A = \cap . A_n$. Then A is a compact subset of a simple link. If now $p \in A$, $T(p) \in A_n$ for all n and hence $T(p) \in A$. Therefore $T(A) \subset A$. Furthermore, if $p \in A$, $T^{-1}(p) \cap \pi$ is contained in every A_n, and hence in A, and $p \in T(A)$. Therefore $T(A) \supset A$ and the theorem is proved.

As a corollary to 4.5 and 2.4 we have

4.8. THEOREM. *If $T: M \to M$ is continuous, and if T carries each simple link into a subset of a simple link—if, for example, the inverse of no point separates a simple link in M—then there exists a simple link F such that $T(F) \subset F$.*

For Peano continua we have

4.9 THEOREM. *If $T: M \to M$ is a continuous transformation of a Peano continuum into itself then*

a) *there exists a subcontinuum π of a cyclic element such that $T(\pi) \supset \pi$.*

b) *either there is a fixed point or some cyclic element intersects its image in a non-degenerate continuum.*

c) *there is a compact subset A of a cyclic element which is fixed under T.*

d) *If T carries each cyclic element into a subset of a cyclic element—if, for example, the inverse of no point separates a cyclic element—then T carries some cyclic element into a subset of itself.*

e) (*Ayres* (see ref. 1)) *if T is a homeomorphism T carries some cyclic element into a subset of itself.*

f) (*Scherrer* (see ref. 6)) *if M is a dendrite then there is a fixed point under T.*

UNIVERSITY OF CHICAGO.

INEQUALITIES OF HIGHER DEGREE IN ONE UNKNOWN.*

By Bruce Elwyn Meserve.

1. Introduction. The determination of the content of linear systems of inequalities has been discussed in detail (see [1] for a bibliography), whereas the corresponding problem for non-linear systems has received little consideration, the object of such a work as [3] being to prove inequalities rather than to solve them. The main purpose of the present paper is to discuss the content of the system which asserts that each member of a finite set of polynomials in a single indeterminate is positive.

A system S composed of a finite number of inequalities of the form

$$0 < f, \tag{1.1}$$

where f is a real-valued function of real variables $x_1, x_2, \cdots, x_m$, has as *content* all sets of real x's (or points) for which (1.1) is satisfied. If f is continuous, the content of (1.1) is an open set composed of regions whose boundaries are contained in the set of real zeros of f.

Here we are concerned only with the case where f is a polynomial in a single indeterminate. The content then consists of a finite number (possibly zero) of non-overlapping (possibly abutting) segments $\alpha < x < \beta$.

The present paper determines (**12**) the number of segments $K(a, b)$ in the content of a finite system

$$0 < g_i(x) \qquad (i = 1, 2, \cdots, s)$$

in $a < x \leqq b$, where a, b are real numbers or $\pm \infty$. It also gives (**8**) bounds for $K(a, b)$, in particular, a bound in terms of $k_1(a, b), \cdots, k_s(a, b)$ which denote the number of segments in the contents of the individual inequalities. There is no restriction on the polynomials, the case of common zeros being handled (**9**) by rational factorization.

The most noteworthy previous paper which deals with the subject and which has come to our attention is by C. F. Gummer [2]. He gives a method for determining what we denote by $K(-\infty, \infty)$ under certain conditions

* Received August 13, 1946; Presented to the American Mathematical Society, August 20, 1946.

[2; 268, 278], the most restrictive of which from the standpoint of inequalities is that the zeros be simple. His process being primarily designed to determine the relative order of the real zeros of the polynomials, his treatment of inequalities is somewhat incidental but clearly brings out the fundamental importance of the Cauchy-Sylvester theorem for the theory. The present paper employs a generalized form of that theorem to give a different process which is both simpler and unrestricted.

An elegant treatment of mixed systems has been given by A. Markoff in *Recueil Mathématique,* vol. 7 (1940), pp. 3-6.

2. Segments and their associated polynomials. Let (α, β) denote a segment (open) and $[\alpha, \beta]$ an interval (closed). For our purposes the general linear inequality can be regarded as having one of the forms

$$0 < x - \alpha \quad \text{or} \quad 0 < -x + \beta \tag{2.1}$$

with the segments (α, ∞), $(-\infty, \beta)$ for respective contents. Similarly, the general quadratic inequality has one of the forms

$$0 < -x^2 + 2bx - c \quad \text{or} \quad 0 < x^2 + 2bx + c \tag{2.2}$$

with $(\alpha, \beta) + (\beta, \alpha)$ and the complement of $[\alpha, \beta] + [\beta, \alpha]$ for respective contents, where α, β are the zeros of the corresponding right member and (α, β), $[\alpha, \beta]$ are to be interpreted as vacuous except when α, β are real numbers satisfying $\alpha < \beta$, $\alpha \leqq \beta$ respectively. From the standpoint of content (2.1) can be regarded as quadratics of the first type in (2.2) if their right members are regarded as having the real zeros α, ∞ and $\beta, -\infty$ respectively.

Conversely, with the exception of the whole line of points, every segment (α, β), where α, β are distinct real numbers or $\pm \infty$ is the content of a unique inequality of one of the three types: $0 < x - \alpha$; $0 < -x + \beta$; $0 < -x^2 + 2bx - c$, with $0 < b^2 - ac$. The segment will be said to *belong to* the corresponding polynomial; the polynomial to *represent* the segment. In general a polynomial of degree n is positive for interior points and zero for boundary points of each of its segments.

3. Content of a polynomial inequality. Let $f(x)$ be a polynomial with real coefficients. The real zeros of f divide the line into a finite number of segments on which f has constant sign. Let $\pi_1, \pi_2, \cdots, \pi_k$ be the poly-

nomials representing the segments on which f is positive. Since the boundaries of the π's are zeros of f, we have

$$f(x) = \sigma\pi_1\pi_2 \cdots \pi_k. \tag{3.1}$$

If $0 < f(x)$, one of the π's is positive, the $k-1$ others are negative and σ has the same sign as $(-1)^{k-1}$. If $0 < -f(x)$, all of the π's are negative and again σ has the sign of $(-1)^{k-1}$. The polynomial σ vanishes at all zeros of f except simple zeros and zero minima of multiplicity exactly two.

The π's can be determined theoretically by finding the real zeros of f. From a knowledge of the real zeros, their multiplicities and the sign of the initial of f, the graphical representation of the segments is rapidly made. A zero x of even multiplicity is a zero minimum if and only if

$$0 < (-1)^{N(x,\infty)} \operatorname{sgn} a_0, \tag{3.2}$$

where $N(x, \infty)$ is the number of real zeros greater than x, each zero being counted a number of times equal to its multiplicity, and a_0 is the initial of f.

4. The number of segments. Our main problem is to determine k, which is the number of segments on which $f(x)$ is positive. A first step in this direction is given by

THEOREM 4.1. *The number k of segments on which $f(x)$ is positive is given by*

$$k = \tfrac{1}{2}\{N_0 + e(-\infty) + e(\infty)\} + \mu,$$

where

$$e(x) = 0, \text{ if } f(x) \leqq 0, \qquad e(x) = 1, \text{ if } 0 < f(x),$$

N_0 is the number of real zeros of $f(x)$ of odd multiplicity, and μ is the number of zero minima of $f(x)$.

The number of segments which belong to f is given by the same formula. Bounds for the number of segments are given by

$$\tfrac{1}{2}\{N_0 + e(-\infty) + e(\infty)\} \leqq k \leqq \tfrac{1}{2}\{N_0 + e(-\infty) + e(\infty)\} + N_e,$$

where N_e is the number of real zeros whose multiplicity is even, and by

$$0 \leqq k \leqq \left[\frac{m}{2}\right] + 1, \tag{4.1}$$

where $m = \deg f$ and the brackets denote "the greatest integer in."

Similar formulas hold for counting the number of segments $k(a, b)$ on which f is positive and which are contained in the point set

$$a < x \leqq b. \tag{4.2}$$

For example,

$$k(a, b) = \tfrac{1}{2}\{N_0(a, b-0) + e(a+0) + e(b-0)\} + \mu(a, b-0), \tag{4.3}$$

where $e(a+0)$ denotes the limit from the right and $e(b-0)$ the limit from the left. Either or both of the extreme segments counted in this formula may fail to *belong to* f because f may not vanish everywhere on the boundary.

Unlike $N_0(a, b)$, the function $k(a, b)$ is not an additive function of (4.2) because an end segment does not necessarily belong to the polynomial f.

The numbers N_0, N_e can be determined by the Sturm sequences [4]. On the other hand, the Sturm sequences for

$$f(x) = (x-1)(x-2)^2,$$

which has a zero minimum, and for

$$g(x) = (x-1)^2(x-2),$$

which does not have a zero minimum, have identical signs at $-\infty$ and at $+\infty$. Thus the zero minima cannot be counted by means of the signs of the Sturm polynomials for f alone. In the general case something like the Cauchy-Sylvester theorem (**6**) seems necessary for the separation of the zero maxima from the zero minima.

5. Number of zero extrema. By rational operations, including differentiation, the polynomial f can be written in the form

$$f = a_0 f_1 f_2^2 \cdots f_r^r, \tag{5.1}$$

where a_0 is the initial of f, each f_i has simple roots and initial 1, and each pair of f_i's is relatively prime. N_0 and N_e may be determined from the Sturm sequences of $f_1, f_2, \cdots, f_r$. N_0 is the total number of real zeros of $f_1, f_3, \cdots$, and N_e is the total number of real zeros of $f_2, f_4, \cdots$.

The zero minima of multiplicity $2h$ are the roots of the *mixed system*

$$0 = g_1, \qquad 0 < g_2, \tag{5.2}$$

where

$$g_1 = f_{2h}, \qquad g_2 = f_1 f_3 \cdots,$$

the zeros of g_2 being the zeros of f with odd multiplicitly. The next section will show how to count the roots of (5.2).

6. The fundamental lemmas. The Cauchy-Sylvester theorem employed by Gummer [2] can be generalized to cover the case of multiple zeros in exactly the same way as Sturm's theorem [4]. Let g_1, g_2 be two relatively prime polynomials with real coefficients (a similar process and result are readily seen to apply when g_1, g_2 have no common real zero on $a < x \leqq b$).

Form the division sequence

$$H_1, H_2, \cdots, H_q, \tag{6.1}$$

in which

$$H_i = H_{i+1}Q_{i+1} - H_{i+2} \qquad (i = 1, 2, \cdots, q-1), \tag{6.2}$$

$$H_1 = g_1, \qquad H_2 = g'_1 g_2, \qquad H_{q+1} = 0.$$

Note that in general H_2 is of higher degree than H_1, so that H_3 is usually $-g_1$.

First suppose that H_q is a non-zero constant. Relations (6.2) can be used to prove

(i) The number of variations of sign in (6.1) can change only at a zero of H_1.

This result can be applied to the subsequence obtained by omitting the first term of (6.1) because that subsequence is generated by applying the division process to H_2, H_3 which are relatively prime in the case being considered. Hence the number of variations of sign in the subsequence can change only at a zero of H_2. Since no zero of H_1 is a zero of H_2, we have

(ii) The number of variations of sign in

$$H_2, H_3, \cdots, H_q$$

does not change at a zero of g_1.

As x increases through a zero of g_1, the polynomial g_1^2 decreases to zero and then increases. Hence $g_1 g'_1$ is negative just before a zero of g_1 and positive just after. Therefore $H_1 H_2$ has opposite sign to g_2 just before the zero and the same sign as g_2 just after. Hence

(iii) Sequence (6.1) loses one variation of sign at a zero of g_1 if g_2 is positive and gains one variation of sign just after a zero of g_1 if g_2 is negative.

Let $F(x)$ be the number of variations of sign in (6.1) for the real number x. On $a < x \leqq b$ let P, N be the number of zeros of g_1 at which g_2 is respectively positive and negative. Then

$$P - N = F(a) - F(b + 0).$$

Except when $g_2(b) < 0 = g_1(b)$, $F(b + 0) = F(b)$.

Now suppose that H_q is not necessarily constant. Divide (6.1) by H_q to give

$$K_1, K_2, \cdots, 1, \tag{6.3}$$

a sequence which can be generated from K_1, K_2 in exactly the same way as (6.1) was generated from H_1, H_2. The zeros of K_1 are simple and are the distinct zeros of g_1. From $H_1H_2 = H_q^2K_1K_2$ we see that the behavior of K_1K_2 near a zero of g_1 is the same as that of H_1H_2. Denote the number of variations of sign in the sequence (6.3) by $F_{11}(x)$.

If $H_q(x)$ is not constant, form the division sequence for H_q, H'_qg_2, remove the highest common factor from its terms and denote the number of variations by $F_{12}(x)$. The zeros of the first term are simple and are the zeros of g_1 with multiplicity at least two.

If the highest common factor of H_q, H'_qg_2 is not constant, form a third sequence with $F_{13}(x)$ variations; and so on.

Put

$$G_{1j} = F_{1j}(a) - F_{1,j+1}(a) - F_{1j}(b + 0) + F_{1,j+1}(b + 0).$$

By taking $g_2 = 1$, we find Sturm sequences for g_1. Denote by S_{1j} the functions corresponding to G_{1j}, the argument $b + 0$ being replaced by b.

THEOREM 6.1. *If g_1, g_2 are relatively prime polynomials with real coefficients, the numbers P_{1j}, N_{1j} of zeros of g_1 which lie on $a < x \leqq b$, which have multiplicity j and which make g_2 respectively positive and negative satisfy*

$$P_{1j} + N_{1j} = S_{1j}, \qquad P_{1j} - N_{1j} = G_{1j}.$$

Theorem 6.1 and the corresponding generalized Sturm theorem which it implies will be called the fundamental lemmas. From them, Theorem 4.1, and (4.3) we see that *the number of non-overlapping segments on which $0 < f(x)$ can be determined by rational operations.*

7. The resolvent. Consider now a finite system S

$$0 < g_i \qquad (i = 1, 2, \cdots, s), \tag{7.1}$$

where the g's are polynomials with real coefficients and a single indeterminate x. Since the intersection of two open sets is an open set, the content of S is a finite number of non-overlapping segments. The system S has the same content as a system containing a single inequality $0 < G$, which will be called a *resolvent* of S.

Various graphical methods for finding the segments of G from those for the g's can be devised. One such is to represent the number x by a point on a semi-circular arc at θ radians from the mid-point of the arc, where $x = \tan\theta$. Erase the segments belonging to $-g_1, -g_2, \cdots, -g_s$ and the boundary points of $g_1, g_2, \cdots, g_s$. The remaining segments are the content of S.

The number of segments which belong to $a < x \leqq b$ and on which G is positive will be denoted by $K(a, b)$. The number of segments on which each g_i is positive will be denoted by $k_i(a, b)$. Usually the (a, b) will be omitted from this notation, but it will nevertheless be understood.

A segment belongs to S if and only if it belongs to a resolvent of S, that is, if and only if every polynomial of S is positive on the segment and at least one polynomial of S vanishes at each point of the boundary of the segment. The phrase a "segment of S," on the other hand, means simply a segment on which G is positive.

8. Bounds on the number of segments. The boundary of the resolvent of S is a subset of the sum of the boundaries of the segments of the polynomials in S. This is merely a restatement of the fact that the real zeros of G are found among those of the product

$$\prod_{i=1}^{s} g_i.$$

Hence by (4.1) the number $K(-\infty, \infty)$ of segments on which G is positive satisfies

$$K(-\infty, \infty) \leqq \left[\tfrac{1}{2}\sum_{i=1}^{s} m_i\right] + 1, \tag{8.1}$$

where $m_i = \deg g_i$.

It is convenient for the present count to regard an infinite segment as having two end points so that the number of end points is always double the number of segments. A left end point of g_1 will be an end point of the system $0 < g_1, 0 < g_2$ if and only if it is either a left end point or an interior point of a segment of g_2. Denote by i_1 the number of end points of g_1 which are interior points of segments of g_2, by i_2 the number of end points of g_2 which are interior points of segments of g_1, by λ the number of points which

are left end points for both g_1 and g_2, and by ρ the number of points which are right ends points for both g_1 and g_2. The total number of end points for $0 < g_1$, $0 < g_2$ is then $i_1 + i_2 + \lambda + \rho$ and the number K of segments is given by

$$2K = i_1 + i_2 + \lambda + \rho. \tag{8.2}$$

If two end points coincide, by shifting one of them slightly we replace the system by another with the same K, k_1, k_2 but with the corresponding pair of end points distinct. Hence in determining formulas for K we may assume that no end points coincide. Similarly, we may arrange that all end points are finite. For the modified system we have

$$2K = i_1 + i_2,$$

$$2k_1 = i_1 + j_1, \qquad 2k_2 = i_2 + j_2,$$

where j_1 is the number of end points of g_1 which are exterior to the segments of g_2 and j_2 is analogously defined. Accordingly,

$$K = k_1 + k_2 - \tfrac{1}{2}(j_1 + j_2).$$

Now the end point farthest to the left is exterior to the segments of g_1 or to the segments of g_2. Similarly for the end point farthest to the right. Hence the least value of $j_1 + j_2$ is 2 and

$$K \leqq k_1 + k_2 - 1.$$

The bound can be attained for all positive integral values of k_1, k_2. This is obvious for $k_1 = k_2 = 1$. Assume the result for k_1, k_2 and add a segment which covers one of the extreme end points. Both K and k_1 (say) are increased by 1 so that the relation

$$K = k_1 + k_2 - 1$$

continues to hold. By $s - 1$ applications of the above we deduce

THEOREM 8.1. *A finite system S (7.1) has K segments, where*

$$K \leqq \sum_{i=1}^{s} k_i - s + 1 \tag{8.3}$$

and k_i denotes the number of segments of g_i. This bound can be attained.

9. Reduction of degree. Two systems having the same content are *equivalent.*

The *degree* of a system of inequalities is the sum of the degrees of the polynomials appearing in the system. Equivalent systems do not necessarily have the same degree.

Let a highest common factor of f, g be h and set

$$f = f^*h, \qquad g = g^*h.$$

Let $S^{(1)}$, $S^{(2)}$ be the systems found by replacing in S the two inequalties $0 < f$, $0 < g$ by the three inequalities

$$0 < f^*, \qquad 0 < g^*, \qquad 0 < h$$

and by the three inequalities

$$0 < -f^*, \qquad 0 < -g^*, \qquad 0 < -h$$

respectively. We may then write

$$S = S^{(1)} S^{(2)}$$

in the sense that every solution of $S^{(1)}$ or $S^{(2)}$ is a solution of S and every solution of S is a solution of $S^{(1)}$ or a solution of $S^{(2)}$. Moreover, $S^{(1)}$ has no solution in common with $S^{(2)}$, that is, no segment of $S^{(1)}$ intersects a segment of $S^{(2)}$. Consequently $K = K_1 + K_2$, where K_i denotes the number of segments of $S^{(i)}$.

Put

$$\deg f^* = m^*, \qquad \deg g^* = n^*, \qquad \deg h = p.$$

Then

$$\deg S = m^* + n^* + 2p + q,$$

$$\deg S^{(1)} = \deg S^{(2)} = m^* + n^* + p + q.$$

Hence, if p is positive,

$$\deg S^{(1)} < \deg S, \qquad \deg S^{(2)} < \deg S.$$

Since the degree of $S^{(i)}$ is finite and non-negative, we have by repetition of the above process

THEOREM 9. 1. *A finite system S may be reduced by rational operations to the product of a finite number of systems such that each pair of polynomials in each factor is relatively prime and such that the number of segments of S equals the sum of the numbers of segments of the factors.*

10. The general system in one indeterminate. In connection with the system S

$$0 < g_1, \quad 0 < g_2, \quad \cdots, \quad 0 < g_s$$

we shall have occasion to consider the *mixed systems* S_i got from S by replacing one inequality $0 < g_i$ by the corresponding equation $0 = g_i$. A number satisfying S_i will be called a *root* of S_i and endowed with the multiplicity which it has as a zero of g_i. In the sections immediately following we shall be interested in the number of roots on a point set $a < x \leqq b$. In general, however, the (a, b) will be omitted from the notation just as $K(a, b)$ has already been abbreviated to K.

Denote the number of real roots of S_i by P_i and the number of real roots of multiplicity j by P_{ij} so that

$$P_i = \sum_{j=1}^{r_i} P_{ij}.$$

Denote by $\mu_{i,2h}$, $M_{i,2h}$ the numbers of zero minima and zero maxima of S_i with multiplicity $2h$. The total numbers of zero extrema are then given by

$$\mu_i = \sum_{h=1}^{\kappa_i} \mu_{i,2h}, \qquad M_i = \sum_{h=1}^{\kappa_i} M_{i,2h},$$

where $2\kappa_i$ is the last even index on a non-vanishing function of the sequence F_{ij} defined in **6**.

Define also the number d_i as the number of roots of S_i got by counting each root of odd multiplicity once, counting each zero minimum twice and omitting each zero maximum.

11. Mixed systems. Consider now the two mixed systems

$$0 = g_1, \quad 0 < g_2, \qquad P_1;$$

$$0 = g_1, \quad 0 < -g_2, \qquad N_1;$$

the capital letter designating in each case the number of roots of a given multiplicity whose index j is omitted for convenience. If we set

$$P_1 + N_1 = S_1, \qquad P_1 - N_1 = G_1, \tag{11.1}$$

the calculation of P_1, N_1 is equivalent to that of S_1, G_1. The number S_1 is really S_{1j} and is not to be confused with the system S_1.

More generally, consider the systems

$$0 = g_1, \quad 0 < g_2, \quad 0 < g_3, \quad P_1;$$
$$0 = g_1, \quad 0 < g_2, \quad 0 < -g_3, \quad Q_1;$$
$$0 = g_1, \quad 0 < -g_2, \quad 0 < g_3, \quad Q_2;$$
$$0 = g_1, \quad 0 < -g_2, \quad 0 < -g_3, \quad Q_3.$$

Then we have also the systems

$$0 = g_1, \quad 0 < g_2, \quad P_1 + Q_1;$$
$$0 = g_1, \quad 0 < -g_2, \quad Q_2 + Q_3;$$
$$0 = g_1, \quad 0 < g_2 g_3, \quad P_1 + Q_3;$$
$$0 = g_1, \quad 0 < g_3, \quad P_1 + Q_2.$$

If we assume that the number of roots can be determined for every fixed system with just one inequality, the numbers $P_1 + Q_1$, $Q_2 + Q_3$, $P_1 + Q_3$, $P_1 + Q_2$ can be calculated and from them P_1, Q_1, Q_2, Q_3.

The same remarks apply to the systems got by adjoining $s - 2$ inequalities to each system above. The general system

$$0 = g_1, \quad 0 < g_2, \quad \cdots, \quad 0 < g_s, \quad P_1$$

for $2 < s$ is one of 2^{s-1} mixed systems, all of which contain $0 = g_1$ and which can be grouped in sets of four. The number of their roots can be computed in terms of the numbers for mixed systems with $s - 2$ inequalities. Hence by induction the calculation of P_1 for the general system is referred back to the numbers S_1 and G_1 appearing in (11.1).

12. The number of segments of a finite system. The numbers in (11.1) and consequently all the numbers in **11** can be determined by the fundamental lemmas provided only that every pair of g's is relatively prime. The factorization of **9** guarantees this for the factors $S^{(1)}, S^{(2)}, \cdots$ of S_1 and the numbers for S_1 can be got by adding those for $S^{(1)}, S^{(2)}, \cdots$.

The $\mu_{1,2h}$ zero minima of S_1 with multiplicity $2h$ are the roots of multiplicity $2h$ of the mixed system

$$0 = g_1, 0 < g_1^{(2h)}, 0 < g_2, \cdots, 0 < g_s.$$

The $M_{1,2h}$ zero maxima of S_1 with multiplicity $2h$ are likewise the roots of multiplicity $2h$ of

$$0 = g_1, 0 < -g_1^{(2h)}, 0 < g_2, \cdots, 0 < g_s.$$

The fundamental lemmas, not directly applicable to these because g_1 and $g_1^{(2h)}$ are not relatively prime, become applicable by the process of **5**.

If g_1, g_2 are two relatively prime polynomials, no segment of g_1 has an end point in common with a segment of g_2. The numbers i_1, i_2 in (8.2) are respectively the d_1, d_2 of **10**. Nothing is contributed to λ, ρ by coincident zeros in this case, but either λ or ρ may be unity because of end points *induced by* the end points of the fundamental interval (a, b). Using the notation of **4** we find the formula

$$K = \tfrac{1}{2}\{d_1(a, b-0) + d_2(a, b-0) + e_1(a+0)e_2(a+0) + e_1(b-0)e_2(b-$$

This formula is easily generalized to

$$K(a, b) = \tfrac{1}{2}\{\sum_{i=1}^{s} d_i(a, b-0) + \prod_{i=1}^{s} e_i(a+0) + \prod_{i=1}^{s} e_i(b-0)\}.$$

As in the case $s = 1$, the number $K(a, b)$ can be found by rational operations.

13. A sufficient condition for consistency in terms of the resultant. If the roots of f are ξ_i, those of g are η_j, $\deg f = m$ and $\deg g = n$, the resultant for the order f, g may be defined as

$$R = (-1)^{\frac{1}{2}m(m-1)} a_0^n \prod g(\xi) = (-1)^{\frac{1}{2}m(m-1)+mn} b_0^m \prod f(\eta).$$

If all roots of f are real, then

$$0 < (-1)^{m+1} \prod g(\xi)$$

is a sufficient condition for one of the numbers $g(\xi)$ to be positive. Hence when all the roots of f are real,

$$0 < (-1)^{\frac{1}{2}m(m+1)+1} a_0^n R, \tag{13.1}$$

where R is expressed in terms of the coefficients of f, g, is a sufficient condition for the consistency of the system

$$0 < f, \qquad 0 < g, \tag{13.2}$$

and when all the roots of g are real

$$0 < (-1)^{\frac{1}{2}m(m-1)+mn+n+1} b_0^m R \tag{13.3}$$

is a sufficient condition. From subsequent developments (Theorem 14.1) it will be apparent that the condition is by no means necessary.

14. The case of two quadratics. Let us examine the condition of consistency for a system of two inequalities, each of which defines a finite segment, that is, in (13.2) suppose

$$f = -x^2 + 2bx - c, \qquad g = -x^2 + 2b'x - c'.$$

Denote the respective segments by (α, β), (γ, δ). They have a point in common if and only if

$$0 < (\alpha - \delta)(\gamma - \beta). \tag{14.1}$$

Multiplication gives

$$0 < \alpha\gamma + \beta\delta - \alpha\beta - \gamma\delta.$$

Setting

$$\alpha = b - d, \qquad \beta = b + d, \qquad \gamma = b' - d', \qquad \delta = b' + d',$$

where

$$0 < d = (b^2 - c)^{\frac{1}{2}}, \qquad 0 < d' = (b'^2 - c')^{\frac{1}{2}},$$

we find

$$\alpha\beta = c, \qquad \gamma\delta = c', \qquad \alpha\gamma + \beta\delta = 2(bb' + dd').$$

Hence condition (14.1) can be written

$$0 < 2(bb' + dd') - c - c'. \tag{14.2}$$

Let us next obtain an equivalent condition by a method which is applicable when all the zeros are real. We have in the present case for the polynomial whose zeros are the values assumed by g at the zeros of f

$$E(0, g) = -g^2 + 2\{c - c' - 2b(b - b')\}g + R,$$

where R is the resultant

$$R = -(c - c')^2 + 4b(b - b')(c - c') - 4c(b - b')^2.$$

The condition that

$$-x^2 + 2Bx - C \qquad (0 \leqq B^2 - C)$$

have at least one positive root is $0 < B$ or $0 < -C$. The condition wanted is that at least one of $E(f, 0)$, $E(0, g)$ have at least one positive root. Hence the condition is that one of the following be satisfied:

$$0 < c - c' - 2b(b - b'), \qquad 0 < c' - c - 2b'(b' - b), \qquad 0 < R.$$

Thus we have

THEOREM 14.1. *The system of inequalities*

$$0 < -x^2 + 2bx - c = f, \qquad 0 < -x^2 + 2b'x - c' = g$$

whose segments are non-vacuous and have no boundary point in common is consistent if and only if

$$T : 0 < 2(bb' + dd') - c - c',$$

where $0 < d = (b^2 - c)^{\frac{1}{2}}$, $0 < d' = (b'^2 - c')^{\frac{1}{2}}$. *This system* T *found by irrational operations may be factored into the product of three systems found by rational operations* $T = UVW$, *where*

$$U : 0 < c - c' - 2b(b - b'), \qquad V : 0 < c' - c - 2b'(b' - b),$$

$$W : 0 < R(f, g).$$

DUKE UNIVERSITY.

REFERENCES.

[1] Lloyd L. Dines and N. H. McCoy, "On linear inequalities," *Transactions of the Royal Society of Canada*, (3) vol. 27 (1933), section III, pp. 37-70.

[2] C. F. Gummer, "The relative distribution of the real roots of a system of polynomials," *Transactions of the American Mathematical Society*, vol. 23 (1922), pp. 265-282.

[3] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge, 1934.

[4] Joseph Miller Thomas, "Sturm's theorem for multiple roots," *National Mathematics Magazine*, vol. 15 (1941), pp. 391-394.

ON RIESZ SUMMABILITY AND SUMMABILITY BY DIRICHLET'S SERIES.* [1]

By C. T. RAJAGOPAL.

We know that, if $k > 0$, summability of any function by the k-th Riesz mean necessarily implies its summability by the Laplace transformation. The converse which is not universally true has been discussed by Szász for the case $k = 1$ [**5**, Theorem 1].[2] Theorem 1 of this paper is a generalization of Szász's problem for $k \geq 1$. Theorem 2 is an extension of a Tauberian oscillation theorem of V. Ramaswami [**3**], involving Riesz means and the Laplace transformation in place of his Cesàro means and Abel mean. Theorems t, T are in the nature of corollaries to Theorems 1, 2.

1. THEOREM 1. *Suppose that*

$$F(t) = t\int_0^\infty A(u)e^{-ut}du \text{ converges for } t > 0, \tag{1}$$

$$F(t) \to s \text{ as } t \to +0; \tag{2}$$

$A_r(x)$ *defined by*

$$A_r(x) = r\int_0^x (x-u)^{r-1}A(u)du, \quad r > 0,$$

$$A_0(x) = A(x),$$

satisfies the condition

$$B_k(x) \equiv (k+1)\{xA_k(x) - A_{k+1}(x)\} \geq -Kx^{k+1} \qquad (x > 0), \tag{3}$$

$k \geq 0$, $K > 0$ *being constants. Then*

$$\sigma_{k+1}(x) \equiv \frac{A_{k+1}(x)}{x^{k+1}} \to s \text{ as } x \to \infty, \tag{4}$$

but $\sigma_k(x)$ *does not in general tend to a limit.*

---

* Received August 2, 1946.

[1] My thanks are due to Dr. V. Ganapathy Iyer and the referee who have suggested the revision of certain details.

[2] Numbers in heavy type refer to the list of references at the end of this paper.

1.1. The proof of the theorem requires the following lemmas of which the first and the second are well known and the third is proved by Szász in the course of his demonstration of the case $k = 0$ of the theorem.

LEMMA 1. *$F(t)$ defined by (1) satisfies the identity:*

$$F(t) = \frac{t^{k+1}}{\Gamma(k+1)} \int_0^\infty A_k(u) e^{-ut} du, \qquad k > 0.$$

LEMMA 2. *If $s(x) \geq 0$ and*

$$t \int_0^\infty s(u) e^{-ut} du \to K \text{ as } t \to +0,$$

then

$$\int_0^x s(u) du \sim Kx \text{ as } x \to \infty.$$

This lemma is a special case of a theorem of Karamata. [**6**, Satz 2].

LEMMA 3. *If*

$$\int_0^x (v(u)/u) du = o(x) \text{ as } x \to \infty, \tag{5}$$

$$t \int_0^\infty \{ \int_0^u (v(x)/x^2) dx \} e^{-ut} du \to s \text{ as } t \to +0, \tag{6}$$

then

$$\int_0^\infty (v(u)/u^2) du = s. \tag{7}$$

LEMMA 4. *Suppositions (1) and (2) imply that*

$$F_k(t) = t \int_0^\infty \sigma_k(u) e^{-ut} du, \qquad k > 0,$$

converges for $t > 0$ and

$$F_k(t) \to s \text{ as } t \to +0. \tag{8}$$

Proof. Lemma 1 gives, when $t > 0$,

$$\begin{aligned} \Gamma(k+1) \int_t^\infty \frac{F(x)}{x^{k+1}} (x-t)^{k-1} dx &= \int_t^\infty (x-t)^{k-1} dx \int_0^\infty A_k(u) e^{-ux} du \\ &= \int_0^\infty A_k(u) du \int_t^\infty (x-t)^{k-1} e^{-ux} dx \\ &= \Gamma(k) \int_0^\infty \frac{A_k(u)}{u^k} e^{-ut} du, \end{aligned}$$

the inversion of the order of integration being justified by absolute convergence. Now

$$F(t)-F_k(t)=kt\int_t^\infty [F(t)-F(x)]\frac{(x-t)^{k-1}}{x^{k+1}}\,dx=kt\int_t^{t^{\frac{1}{2}}}+kt\int_{t^{\frac{1}{2}}}^\infty, \qquad t<1,$$

where, as $t\to 0$,

$$\begin{aligned} kt\int_{t^{\frac{1}{2}}}^\infty F(x)\,\frac{(x-t)^{k-1}}{x^{k+1}}\,dx &= O\Big[kt\int_{t^{\frac{1}{2}}}^\infty \frac{(x-t)^{k-1}}{x^{k+1}}\,dx\Big] \\ &= O\Big[k\int_{x=t^{\frac{1}{2}}}^\infty (1-(t/x))^{k-1}d(1-(t/x))\Big] \\ &= O[1-(1-t^{\frac{1}{2}})^k]. \end{aligned}$$

Similarly

$$kt\int_t^{t^{\frac{1}{2}}}\frac{(x-t)^{k-1}}{x^{k+1}}\,dx=(1-t^{\frac{1}{2}})^k$$

and therefore

$$\begin{aligned} |F(t)-F_k(t)| &< \max_{t\le x\le t^{\frac{1}{2}}}|F(t)-F(x)|(1-t^{\frac{1}{2}})^k \\ &\quad +[1-(1-t^{\frac{1}{2}})^k][|F(t)|+O(1)], \qquad t\to 0, \end{aligned}$$

which, in conjunction with (2), establishes (8).

1.2. Proof of Theorem 1. Lemma 4 for $F_k(t)$ and $F_{k+1}(t)$ can be used in (3) so as to obtain successively

$$t\int_0^\infty \frac{B_k(u)}{u^{k+1}}\,e^{-ut}du\to 0 \text{ as } t\to 0,$$

$$t\int_0^\infty \{\frac{B_k(u)}{u^{k+1}}+K\}e^{-ut}du\to K \text{ as } t\to 0.$$

Hence we can take, in Lemma 2,

$$s(u)=\frac{B_k(u)}{u^{k+1}}+K$$

and get

$$\int_0^x \{\frac{B_k(u)}{u^{k+1}}+K\}du\sim Kx \text{ as } x\to\infty,$$

or

$$\int_0^x \frac{B_k(u)}{u^{k+1}}\,du=o(x) \text{ as } x\to\infty. \tag{9}$$

Also we have, from the definition of $B_k(x)$ in (3),

$$B_k(x) = x^{k+2}\frac{d}{dx}\frac{A_{k+1}(x)}{x^{k+1}}$$

which gives rise to

$$\sigma_{k+1}(x) = \int_0^x \frac{B_k(u)}{u^{k+1}}\,du. \tag{10}$$

When we write $(k+1)$ in place of k in (8) and use (10), we have a relation which is (6) with $v(u) = B_k(u)/u^k$. The same choice of $v(u)$ makes (5) identical with (9). Lemma 3 therefore enables us to conclude that

$$\int_0^\infty \frac{B_k(u)}{u^{k+2}}\,du = s, \text{ i. e., } \sigma_{k+1}(x) \to s \text{ as } x \to \infty.$$

A necessary condition for the further conclusion $\sigma_k(x) \to s$ is $B_{k+1}(x)/x^{k+1} \to 0$ which is more than (3). Hence the further conclusion is not in general warranted.

1. 3. In Theorem 1, we can take $A(x) = \sum_{\lambda_n \leqq x} a_n$, where $\{\lambda_n\}$ is a positive, increasing, divergent sequence, and obtain

COROLLARY 1. *Suppose that*

$$F(t) = \sum_{\nu=1}^{\infty} a_\nu e^{-\lambda_\nu t}, \qquad 0 < \lambda_1 < \lambda_2 < \cdots, \lambda_n \to \infty;$$

$$A_r(x) = \sum_{\lambda_n \leqq x} (x - \lambda_n)^r a_n,$$

$$\frac{B_r(x)}{r+1} = \sum_{\lambda_n \leqq x} (x - \lambda_n)^r a_n \lambda_n, \qquad r \geqq 0,$$

Then the conditions:

(i) $F(t)$ *converges for* $t > 0$, $\lim_{t \to +0} F(t) = s$,

(ii) $B_k(x) \geq -k\lambda_n^{k+1}, \qquad \lambda_n \leq x < \lambda_{n+1}, \qquad k \geq 0,$

together imply that $\sum_{\nu=1}^{\infty} a_\nu$ *is summable-*$R(\lambda_n, k+1)$, *i. e., summable by the Riesz mean of type* λ *and order* $(k+1)$; *but not in general summable-*$R(\lambda_n, k)$.

This result leads at once to the following extension of a classical theorem of Schnee [**2**, Theorem 33].

COROLLARY 2. *Necessary and sufficient conditions for the* $R(\lambda_n, k)$*-summability* $(k \geq 0)$ *of* $\sum_{\nu=1}^{\infty} a_\nu$ *are* (*in the notation of Corollary* 1):

(i) $F(t)$ *converges for* $t>0$, $\lim_{t\to+0} F(t)=s$,

(ii′) $B_k(x)=o(\lambda_n^{k+1})$, $\lambda_n\leq x<\lambda_{n+1}$ $\quad(n\to\infty)$.

That the conditions are necessary for the $R(\lambda_n,k)$-summability of $\sum_{\nu=1}^{\infty}a_\nu$ follows from Lemma 1 and the definition of $B_k(x)$.

To prove the sufficiency of the conditions we observe, from Corollary 1, that (i) and (ii′) establish the $R(\lambda_n,k+1)$-summability of $\sum_{\nu=1}^{\infty}a_\nu$. The $R(\lambda_n,k)$-summability of the series follows as a consequence of (ii′).

2. When $\lim_{t\to+0} F(t)$ does not exist, we can (under certain conditions) determine an upper or a lower bound for $\overline{\lim_{t\to\infty}}\,\sigma_k(x)$, for large enough k, in terms of $\overline{\lim_{t\to+0}}\,F(t)$ or $\underline{\lim_{t\to+0}}\,F(t)$. We have, in this direction,

LEMMA 5. *Let us write*

$$\overline{\lim_{t\to+0}}\,F(t)=\bar{s},\qquad \overline{\lim_{x\to\infty}}\,\sigma_k(x)=\bar{\sigma}_k.$$

Let us supose that $r\geq 0$, $k\geq 1$ *and write*

$$\nu_k^{(r)}=\frac{\Gamma(k)}{\Gamma(r+k+1)}\,\frac{(r+k)^{r+k}}{(k-1)^{k-1}}\,.^{3}$$

Then, if $A_r(x)\geq 0$ $(x\geq 0)$,

(11) $$\bar{\sigma}_{r+k}\leq\frac{e^{r+1}}{\nu_k^{(r)}}\,\bar{s}$$

provided the right-hand side is finite.

Proof. From the known result [2, Lemma 6]:

(12) $$A_{r+k}(x)=\frac{\Gamma(r+k+1)}{\Gamma(k)\Gamma(r+1)}\int_0^x(x-u)^{k-1}A_r(u)\,du,$$

we obtain

(13) $$\sigma_{r+k}(x)=\frac{\Gamma(r+k+1)}{\Gamma(k)\Gamma(r+1)}\,y^{r+1}\int_0^\infty g(u)\,e^{-uy}A_r(u)\,du,$$

where

$$g(u)=\frac{(x-u)^{k-1}e^{uy}}{x^{r+k}y^{r+1}}$$

reaches its maximum when $u=x-\dfrac{k-1}{y}$, the maximum being

[3] When $k=1$, $(k-1)^{k-1}$ in the expression for $\nu_k^{(r)}$ is to mean 1.

$$\frac{e^{xy}}{(xy)^{k+r}} e^{-(k-1)} (k-1)^{k-1}.$$

Since this maximum is least when $xy = k + r$, (13) gives, with the last-mentioned relation between x and y,

$$\sigma_{r+k}(x) \leq \frac{e^{r+1}}{\nu_k^{(r)}} \frac{1}{\Gamma(r+1)} y^{r+1} \int_0^x e^{-uy} du \int_0^u r(u-t)^{r-1} A(t)\, dt.^4$$

In this inequality we can invert the order of integration and let $x \to \infty$ (or $y \to +0$), obtaining

$$\begin{aligned} \bar{\sigma}_{r+k} &\leq \frac{e^{r+1}}{\nu_k^{(r)}} \frac{1}{\Gamma(r+1)} \overline{\lim_{y \to +0}} y^{r+1} \int_0^\infty A(t)\, dt \int_t^\infty r(u-t)^{r-1} e^{-uy} du \\ &= \frac{e^{r+1}}{\nu_k^{(r)}} \frac{1}{\Gamma(r+1)} \overline{\lim_{y \to +0}} y \int_0^\infty A(t) e^{-ty} \Gamma(r+1)\, dt \end{aligned}$$

which is the conclusion sought.

A particular version of Lemma 5, involving the $(r+k)$-th Cesàro mean and the Abel mean, has been given by Garten and Knopp [**1**, Satz 4]. The lemma can be used in much the same way as its particular version and made to yield the next two results.

LEMMA 6. *If (in the notation of Lemma 5), for some $r \geq 0$, $\sigma_r(x) \geqq -K$, then*

$$\bar{\sigma}_\infty \equiv \lim_{k \to \infty} \bar{\sigma}_k = \bar{s}.$$

Proof. Starting from the following relation, equivalent to (12),

$$A_{r+k}(x) + Kx^{r+k} = \frac{\Gamma(r+k+1)}{\Gamma(k)\Gamma(r+1)} \int_0^x (x-u)^{k-1} \{A_r(u) + Ku^r\} du,$$

we can get, arguing as in Lemma 5,

$$\bar{\sigma}_{r+k} + K \leq \frac{e^{r+1}}{\nu_k^{(r)}} (\bar{s} + K)$$

whence, letting $k \to \infty$, it follows that

$$\bar{\sigma}_\infty \leq \bar{s}$$

which, in conjunction with the universal relation $\bar{s} \leq \bar{\sigma}_\infty$, establishes the conclusion of the theorem.

The above argument tacitly asumes that $\bar{s} < \infty$, the case $\bar{s} = \infty$ being trivial.

---

[4] From this point onwards, the proof tacitly assumes $r > 0$, the argument in the case $r = 0$ being an obvious simplification of what follows.

2. 1. Lemma 6 can be stated in a more comprehensive form as follows.

THEOREM 2. *Let*

$$\overline{\lim_{t\to+0}} F(t) = \bar{s}, \quad \overline{\lim_{x\to\infty}} \sigma_k(x) = \bar{\sigma}_k, \quad \lim_{k\to\infty} \bar{\sigma}_k = \bar{\sigma}_\infty.$$

Then (a) *provided* $\bar{\sigma}_\infty$ *is finite, and consequently, for large enough* r, $\sigma_r(x)$ *is bounded below,*

$$\bar{\sigma}_\infty = \bar{s};$$

(b) *provided* $\bar{\sigma}_\infty$ *is finite,*

$$\underline{\sigma}_\infty = \underline{s};$$

(c) *provided* $\underline{\sigma}_\infty$ *and* $\bar{\sigma}_\infty$ *are both finite,*

$$\underline{\sigma}_\infty = \underline{s}, \qquad \bar{\sigma}_\infty = \bar{s}.$$

3. Concluding remarks. Suppose that $\bar{\sigma}_\infty$ and $\underline{\sigma}_\infty$ have the common finite value s. Then, for large enough r, $\bar{\sigma}_r$ and $\underline{\sigma}_r$ are finite. And so $\sigma_r(x)$ is bounded for $r \geq k$ (say), which ensures the hypotheses (3) of Theorem 1. Further, the universal fact $\underline{\sigma}_\infty \leq \overline{\lim_{t\to+0}} F(t) \leq \bar{\sigma}_\infty$ implies the hypothesis (2) of Theorem 1, viz., $\lim_{t\to+0} F(t) = s$. Therefore, appealing to Theorem 1, we find that $\bar{\sigma}_r = \underline{\sigma}_r = s$ for $r \geq k+1$. Thus we establish

THEOREM t. *In the notation of Theorem* 2, $\bar{\sigma}_\infty = \underline{\sigma}_\infty = s$ (*finite*) *ensures* $\bar{\sigma}_r = \underline{\sigma}_r = s$ *for all sufficiently large* r.

Next supposing that $\bar{s}$ and $\underline{s}$ have the common finite value s, and further that $\underline{\sigma}_k > -\infty$, we can conclude from Theorem 2 that $\bar{\sigma}_\infty = s = \underline{\sigma}_\infty$. This fact, taken along with Theorem t, proves

THEOREM T. *In the notation of Theorem* 2, $\bar{s} = \underline{s} = s$ (*finite*) *and* $\underline{\sigma}_k > -\infty$ *together ensure* $\bar{\sigma}_r = \underline{\sigma}_r = s$ *for all sufficiencly large* r.

The forms of Theorems 2, t, T involving Cesàro means and the Abel mean are well known [**3**, Theorems 1(a), (b), (c); **4**, Theorems t, T].

If $A(x) = \sum_{\lambda_n \leq x} a_n$ $(0 < \lambda_n < \lambda_{n+1} \to \infty)$ in §§ 2, 3, the propositions given there assume the form of statements connecting the summability of $\sum a_n$ by Riesz means with the summability of the series by the $D(\lambda_n)$-process, i. e., by Dirichlet's series.

MADRAS CHRISTIAN COLLEGE,
TAMBARAM, S. INDIA.

REFERENCES.

1. V. Garten and K. Knopp, "Ungleichungen zwischen Mittelwerten von Zahlenfolgen und Funktionen," *Mathematische Zeitschrift*, vol. 42 (1937), pp. 365-388.
2. G. H. Hardy and M. Riesz, *The general theory of Dirichlet's series* (Cambridge Tracts, No. 18, 1915).
3. V. Ramaswami, "Some Tauberian theorems on oscillations," *Journal of the London Mathematical Society*, vol. 10 (1935), pp. 294-308.
4. J. E. Littlewood, "Note on the preceding paper," *Journal of the London Mathematical Society*, vol. 10 (1935), pp. 309-310.
5. O. Szász, "Converse theorems of summability for Dirichlet's series," *Transactions of the American Mathematical Society*, vol. 39 (1936), pp. 117-130.
6. J. Karamata, "Neuer Beweis und Verallgemeinerung einiger Tauberian-Sätze," *Mathematische Zeitschrift*, vol. 33 (1931), pp. 294-299.

A STUDY ON THE THEORY OF CONJUGATE NETS.*

By CHUAN-CHIH HSIUNG.

1. Introduction. This paper is concerned with the study of the projective differential geometry of curves of a conjugate net on an analytic non-ruled surface in ordinary space. Some portions of the theory of a surface S referred to a conjugate net N_x, which are used in later developments, are summarized in **2**.

Let C_λ be any curve on the surface S through a general point P_x. As the point P_x moves along the curve C_λ, the parametric u-, v-tangents generate respectively two non-developable ruled surfaces $R^{(u)}$, $R^{(v)}$. In **3** we find the correspondence between the line joining any two points T, $\bar{T}$ on the parametric u-, v-tangents through the point P_x and the intersection of the tangent planes of the ruled surfaces $R^{(u)}$, $R^{(v)}$ at the points T, $\bar{T}$ respectively. In **4** two quadrics are introduced of which one has contact of the second order with the ruled surface $R^{(u)}$ at the point T and contact of the first order with the ruled surface $R^{(v)}$ at the point $\bar{T}$, and the other has the same properties with the roles of u, v interchanged.

By *a line* l_1 we mean, as usual, any line through the point P_x of the surface S and not lying in the tangent plane of the surface S at P_x; by *a line* l_2 we mean, dually, any line in the tangent plane of the surface S at P_x but not passing through the point P_x. In **5**, by means of certain correspondences between l_1 and l_2 we construct a one-parameter family of polarities and present a new geometric characterization for a general canonical line (of each kind) of Davis.

In the last section we derive two one-parameter families of lines, of which two may be regarded as generalizations of the canonical edges of Green of the asymptotic net of a surface to a conjugate net.

2. Analytic basis. If the projective homogeneous coordinates $x^{(1)}, \cdots, x^{(4)}$ of a point P_x in ordinary space are given as analytic functions of two independent variables u, v by equations of the form

$$x = x(u, v), \tag{1}$$

* Received September 19, 1946.

the locus of P_x as u, v vary is an analytic surface S. If the parametric curves on the surface form a conjugate net N_x, the four coordinates x and the four coordinates of the point P_y, which is the harmonic conjugate of the point P_x with respect to the foci of the axis of the point P_x, satisfy a system of partial differential equations of the form [7, pp. 250-254]

$$(2)\qquad \begin{aligned} x_{uu} &= px + \alpha x_u + Ly, \\ x_{uv} &= cx + ax_u + bx_v, \\ x_{vv} &= qx + \delta x_v + Ny \qquad (LN \neq 0), \end{aligned}$$

subscripts indicating partial differentiation and the coefficients being functions of u, v which satisfy certain integrability conditions. It is easy to verify that

$$(3)\qquad y_u = fx - nx_u + sx_v + Ay, \qquad y_v = gx + tx_u + nx_v + By,$$

where we have placed

$$(4)\qquad \begin{aligned} fN &= c_v + ac + bq - c\delta - q_u, & gL &= c_u + bc + ap - c\alpha - p_v, \\ -nN &= a_v + a^2 - a\delta - q, & tL &= a_u + ab + c - \alpha_v, \\ sN &= b_v + ab + c - \delta_u, & nL &= b_u + b^2 - b\alpha - p, \\ A &= b - (\log N)_u, & B &= a - (\log L)_v. \end{aligned}$$

The ray-points, or Laplace transformed points, ρ, σ of the point P_x are given by the formulas

$$(5)\qquad \rho = x_u - bx, \qquad \sigma = x_v - ax.$$

Some of the invariants of the parametric conjugate net are

$$(6)\qquad \begin{aligned} H &= c + ab - a_u, & K &= c + ab - b_v, \\ \mathfrak{H} &= sN, & \mathfrak{K} &= tL, \\ 8\mathfrak{B}' &= 4a - 2\delta + l_v, & 8\mathfrak{C}' &= 4b - 2\alpha - l_u, \\ r &= N/L, \end{aligned}$$

where l is defined by placing

$$l = \log r.$$

We shall suppose that $\mathfrak{H}\mathfrak{K} \neq 0$ so that the parametric curves are not plane curves.

If the points x, ρ, σ, y are used as the vertices of a local tetrahedron of reference with a unit point suitably chosen, then any point given by an expression of the form

$$x_1 x + x_2 \rho + x_3 \sigma + x_4 y \tag{7}$$

has local coordinates proportional to $x_1, \cdots, x_4$.

Introducing non-homogeneous coordinates by the definitions

$$x = x_2/x_1, \qquad y = x_3/x_1, \qquad z = x_4/x_1, \tag{8}$$

we obtain a power series expansion [6, p. 481] for one non-homogeneous coordinate z of a point on the surface S in terms of the other two coordinates x, y:

$$z = \tfrac{1}{2}(Lx^2 + Ny^2) + \tfrac{1}{3}(L\mathfrak{C}'x^3 + N\mathfrak{B}'y^3) + \cdots. \tag{9}$$

3. A Correspondence. Let us consider a point P_x on a surface S and the curve C_λ passing through P_x and belonging to the family defined on S by the equation

$$dv - \lambda du = 0, \tag{10}$$

where λ is a function of u, v; and let the parametric representation of the curve C_λ be

$$u = u(w), \qquad v = v(w). \tag{11}$$

Without loss of generality we may assume that, for the point P_x, $w = 0$. Throughout this paper, differentiation with respect to w is denoted by accents and $u', u'', \cdots; v', v'', \cdots$ are the derivatives of u, v at any point on the parametric u-tangent through the point P_x, so that

$$\lambda = v'/u', \qquad \lambda' = v''/u'^2 - v'u''/u'^3. \tag{12}$$

The parametric u-, v-tangents at points of the curve C_λ generate two non-developable ruled surfaces $R^{(u)}$, $R^{(v)}$ respectively. The points

$$T = ex - \rho, \qquad \bar{T} = \bar{e}x - \sigma, \tag{13}$$

where $e, \bar{e}$ are functions of u, v, lie on the parametric u-, v-tangents through the point P_x respectively, and the line l_2 determined by them lies in the tangent plane of the surface S at the point P_x.

From system (2) and equations (3), by differentiation and substitution, any derivative of T with respect to u, v can be expressed as a linear combination of x, x_u, x_v, y. In particular, making use of equations (4), (6) one obtains

$$
\begin{aligned}
T_u &= (e_u + b_u - p)x + (e + b - \alpha)x_u - Ly, \\
T_v &= (e_v + b_v - c)x - ax_u + ex_v, \\
T_{uu} &= (*)x + (*)x_u - sLx_v + [L(e + b - \alpha) - L_u - AL]y, \\
T_{uv} &= (*)x + (*)x_u + (e_u + eb)x_v - aLy, \\
T_{vv} &= (*)x + (*)x_u + (2e_v + e\delta - K)x_v + eNy.
\end{aligned} \tag{14}
$$

where (*) denotes terms immaterial for our purpose.

The tangent planes to the ruled surface $R^{(u)}$ at the point T and to the ruled surface $R^{(v)}$ at the point $\bar{T}$ intersect in a line l_1 which passes through the point P_x. There must be a point on the line l_1 given by an expression of the form

$$z = y + \xi\rho + \eta\sigma.$$

The plane of the points x, z, T is to be tangent to the ruled surface $R^{(u)}$ at T, and therefore must contain the tangent to the curve traced by T as the point P_x moves along the curve C_λ. It is clear from the first two of equations (14) that a point on this tangent is given by

$$T_w = (*)x + (*)\rho + ev'\sigma - Lu'y;$$

this is to lie in the plane xzT, which can occur if and only if $\eta = -ev'/Lu'$. Similarly, the plane $xz\bar{T}$ is tangent to the ruled surface $R^{(v)}$ at $\bar{T}$ if and only if $\xi = -\bar{e}u'/Nv'$. Thus we obtain a correspondence between the line l_2 in the tangent plane $x_4 = 0$ and the line l_1 through the point P_x:

$$Lx_3 + e\lambda x_4 = 0, \qquad N\lambda x_2 + \bar{e}x_4 = 0. \tag{15}$$

It is easily seen that *this correspondence becomes a polarity with respect to a quadric (and therefore ∞^1 quadrics) if and only if λ satisfies*

$$L - N\lambda^2 = 0, \tag{16}$$

that is, *if and only if the tangent of the curve C_λ at the point P_x is an associate conjugate tangent of the parametric conjugate net N_x.* The equation of any one of these quadrics is found to be of the form

$$\sqrt{LN}\,x_2x_3 - x_1x_4 + k_4x_4^2 = 0, \tag{17}$$

where k_4 denotes a parameter. If a unique quadric of this pencil is desired, we may choose the one that passes through the covariant point P_y. For this quadric we have $k_4 = 0$.

4. Two quadrics associated with the points T, $\bar{T}$. The parametric vector equation of the ruled surface $R^{(u)}$ is

$$X = T + \mu x, \tag{18}$$

where μ, w are taken as independent parameters. The coordinates of any point $\mathfrak{X}$ near the point T on the ruled surface $R^{(u)}$ can be represented by Taylor's expansion as power series in the increments $\Delta\mu$, Δw corresponding to displacement on $R^{(u)}$ from T to the point $\mathfrak{X}$:

$$\begin{aligned}\mathfrak{X} &= X(\Delta\mu, \Delta w) = X(0,0) + X_\mu(0,0)\Delta\mu + X_w(0,0)\Delta w \\ &+ \tfrac{1}{2}[X_{\mu\mu}(0,0)\Delta\mu^2 + 2X_{\mu w}(0,0)\Delta\mu\Delta w + X_{ww}(0,0)\Delta w^2] + \cdots,\end{aligned} \tag{19}$$

where

$$\begin{aligned} X_\mu(0,0) &= x, \qquad X_w(0,0) = T_u u' + T_v v', \\ X_{\mu\mu}(0,0) &= 0, \qquad X_{\mu w}(0,0) = x_u u' + x_v v', \\ X_{ww}(0,0) &= T_{uu}u'^2 + 2T_{uv}u'v' + T_{vv}v'^2 + T_u u'' + T_v v''. \end{aligned} \tag{20}$$

In accordance with equation (14), all the derivatives of X at the point T may be expressed uniquely as linear combinations of x, ρ, σ, y, so that we may write $\mathfrak{X}$ in the form (7), where the local coordinates $x_1, \cdots, x_4$ of the point $\mathfrak{X}$ are given by the expansions

$$\begin{aligned} x_1 &= e + \Delta\mu + [u'(e_u + eb + nL) + v'(e_v + ea - K)]\Delta w + \cdots, \\ x_2 &= -1 + [u'(e + b - \alpha) - av']\Delta w + \cdots, \\ x_3 &= ev'\Delta w + v'\Delta\mu\Delta w + \tfrac{1}{2}[-sLu'^2 + 2(e_u + eb)u'v' \\ &\qquad + (2e_v + e\delta - K)v'^2 + ev'']\Delta w^2 + \cdots, \\ x_4 &= -Lu'\Delta w + \tfrac{1}{2}\{[L(e + b - \alpha) - L_u - AL]u'^2 \\ &\qquad - 2aLu'v' + eNv'^2 - Lu''\}\Delta w^2 + \cdots. \end{aligned} \tag{21}$$

The equation of any quadric, which passes through the points T, $\bar{T}$ and is tangent to the planes (15) at T, $\bar{T}$ respectively, is found to be

$$\begin{aligned} a_{11}(x_1^2 + 2ex_1x_2 + 2\bar{e}x_1x_3 + e^2x_2^2 + \bar{e}^2x_3^2) + 2a_{14}x_1x_4 + 2a_{23}x_2x_3 \\ + 2a_{24}x_2x_4 + 2a_{34}x_3x_4 + a_{44}x_4^2 = 0, \end{aligned} \tag{22}$$

where the coefficients a_{24}, a_{34} are defined by the following formulas:

$$\begin{aligned} a_{24} &= (d/Lu')(Lu'a_{14} + v'a_{23} - e\bar{e}v'a_{11}), \\ a_{34} &= (e/Nv')(Nv'a_{14} + u'a_{23} - e\bar{e}u'a_{11}). \end{aligned} \tag{23}$$

Imposing on equation (22) the further conditions that it be satisfied by the series (21) identically in the terms of the second degree of $\Delta\mu$, Δw, we obtain easily *the equation of the quadric Q_u which has contact of the second order with the ruled surface $R^{(u)}$ at the point T and contact of the first order with the ruled surface $R^{(v)}$ at the point $\bar{T}$,* namely,

$$(24)\quad 2L(\lambda x_1x_4 - Lx_2x_3) - (2\bar{e}/r\lambda)(L - N\lambda^2)x_3x_4 + \{L(2n\lambda - \lambda^2K/L + \mathfrak{H}/N) + e\lambda[4(\mathfrak{C}' + \lambda\mathfrak{B}') - \tfrac{1}{2}(l_u + \lambda l_v) - (\log\lambda)'] - e(2\bar{e}/N - e\lambda/L)(L - N\lambda^2)\}x_4^2 = 0.$$

There is also *a unique quadric Q_v with second order contact with the ruled surface $R^{(v)}$ at the point $\bar{T}$ and first order contact with the ruled surface $R^{(u)}$ at the point T.* The equation of this quadric Q_v can be obtained in a way similar to the foregoing, or else can be written immediately by applying the substitution

$$(25)\qquad \begin{pmatrix} u & 2 & e & s & n & L & H & \mathfrak{B}' & \lambda & \lambda' \\ v & 3 & \bar{e} & t & -n & N & K & \mathfrak{C}' & 1/\lambda & -\lambda'/\lambda^3 \end{pmatrix}$$

to equation (24); the result is

$$(26)\quad 2N(x_1x_4 - N\lambda x_2x_3) + 2er(L - N\lambda^2)x_2x_4 + \{-N(2n + H/\lambda N - \lambda\mathfrak{R}/L) + (\bar{e}/\lambda)[4(\mathfrak{C}' + \lambda\mathfrak{B}') + \tfrac{1}{2}(l_u + \lambda l_v) + \log\lambda)'] + (\bar{e}/\lambda^2)(2e\lambda/L - \bar{e}/N)(L - N\lambda^2)\}x_4^2 = 0.$$

On the other hand, if π be any plane passing through the line $T\bar{T}$, then the plane sections Γ_u, Γ_v of the ruled surfaces $R^{(u)}$, $R^{(v)}$ made by π have simple points at T, $\bar{T}$ respectively. It should be noted [1] that as the plane π revolves about the line $T\bar{T}$, the locus of the conic in π having contact of the second order with the plane section Γ_u at the point T and contact of the first order with the plane section Γ_v at the point $\bar{T}$ is the quadric Q_u, and a similar argument can be made with u, v interchanged.

Since the quadrics Q_u, Q_v both depend upon the points T, $\bar{T}$, we obtain two two-parameter families of quadrics associated with the point P_x of the curve C_λ. In particular, if the points T, $\bar{T}$ are the ray-points ρ, σ of the point P_x, then $e = \bar{e} = 0$, and the equations of the quadrics Q_u, Q_v become respectively

$$(27)\qquad 2(\lambda x_1x_4 - Lx_2x_3) + (2n\lambda - \lambda^2K/L + \mathfrak{H}/N)x_4^2 = 0,$$

[1] For the proof of this fact for two hypersurfaces and its generalization see paper [4] of the author.

$$2(x_1x_4 - N\lambda x_2x_3) - (2n + H/\lambda N - \lambda\mathfrak{R}/L)x_4^2 = 0. \tag{28}$$

The two quadrics (27), (28) intersect, not only in the parametric tangents, but also in a residual conic whose plane is obtained by eliminating the terms containing x_2x_3 from equations (27), (28), namely,

$$2(L - N\lambda^2)x_1 + [rK\lambda^3 + \lambda(\mathfrak{R} - \mathfrak{H}) - H/\lambda r - 2n(L + N\lambda^2)]x_4 = 0, \tag{29}$$

which is a plane passing through the ray $\rho\sigma$. Moreover, the cone projecting this conic of intersection from the point P_x is similarly found to be given by the equation

$$2(L - N\lambda^2)x_2x_3 - [4n\lambda + (1/N)(H + \mathfrak{H}) - (\lambda^2/L)\ (K + \mathfrak{R})]x_4^2 = 0. \tag{30}$$

5. The canonical lines of Davis. A canonical line $l_1(h)$ of the first kind of Davis [1, p. 18] is the intersection of the two planes

$$Lx + h\mathfrak{C}'z = 0, \qquad Ny + h\mathfrak{B}'z = 0, \tag{31}$$

and a canonical line $l_2(h)$ of the second kind crosses the parametric tangents at the point P_x in the points

$$(-(1/h\mathfrak{C}'), 0, 0) \qquad (0, -(1/h\mathfrak{B}'), 0), \tag{32}$$

where h is a constant.

In order to give a new geometric characterization of a general canonical line of each kind of Davis we derive a one-parameter family of correspondences as follows.

Let $\bar{g}$ be the polar line of any line l_1 with the equation

$$x + ez = 0, \qquad y + \bar{e}z = 0, \tag{33}$$

with respect to any quadric of Darboux,

$$Lx^2 + Ny^2 + z(-2 + 4\mathfrak{C}'x + 4\mathfrak{B}'y + k_4z) = 0, \tag{34}$$

at the point P_x of the surface S, where k_4 is a parameter. If $\bar{g}$ intersects the u-, v-tangents respectively in the points $P_1^{(\infty)}(\bar{x}^{(\infty)}, 0, 0)$ and $\bar{P}_2^{(\infty)}(0, \bar{y}^{(\infty)}, 0)$, then

$$1/\bar{x}^{(\infty)} = 2\mathfrak{C}' - eL, \qquad 1/\bar{y}^{(\infty)} = 2\mathfrak{B}' - \bar{e}N. \tag{35}$$

This correspondence between l_1 and the line joining the points $\bar{P}_1^{(\infty)}$, $\bar{P}_2^{(\infty)}$ is called, for brevity, the correspondence C_∞.

On the other hand, let Γ_1 be the plane section of the surface S made by the plane through l_1 and the u-tangent; then by means of (9) we obtain easily the power series expansion of Γ_1 at P_x:

$$y = -(\bar{e}/2)Lx^2 - \tfrac{4}{3}\bar{e}L\mathfrak{C}'x^3 + \cdots. \tag{36}$$

The pole $\bar{P}_1^{(0)}(\bar{x}^{(0)}, 0, 0)$ of l_1 with respect to any four-point conic of Γ_1 at P_x is given by

$$1/\bar{x}^{(0)} = \tfrac{8}{3}\mathfrak{C}' - eL. \tag{37}$$

For the plane section Γ_2 of the surface S made by the plane through l_1 and the v-tangent, we obtain similarly on the v-tangent the pole $\bar{P}_2^{(0)}(0, \bar{y}^{(0)}, 0)$, where

$$1/\bar{y}^{(0)} = \tfrac{8}{3}\mathfrak{B}' - \bar{e}N. \tag{38}$$

Thus we arrive at a correspondence C_0 between l_1 and the line joining the points $\bar{P}_1^{(0)}$, $\bar{P}_2^{(0)}$.

Now on the u-, v-tangents we take respectively two points $\bar{P}_1^{(k)}$, $\bar{P}_2^{(k)}$ such that the cross ratio of the four points $\bar{P}_m^{(\infty)}$, $\bar{P}_m^{(0)}$, P_x and $\bar{P}_m^{(k)}$,

$$(\bar{P}_m^{(\infty)}\bar{P}_m^{(0)}, P_x\bar{P}_m^{(k)}) = k \qquad (m = 1, 2), \tag{39}$$

k being a constant, and define a one-parameter family of correspondences C_k between l_1 and the join of $\bar{P}_1^{(k)}$, $\bar{P}_2^{(k)}$ by the equation

$$z = [\tfrac{2}{3}\left(\frac{4-3k}{1-k}\right)\mathfrak{C}' - eL]x + [\tfrac{2}{3}\left(\frac{4-3k}{1-k}\right)\mathfrak{B}' - \bar{e}N]y - 1 = 0. \tag{40}$$

It is easily shown that *the correspondence C_k is the polarity with respect to any quadric Q_k of the pencil,*

$$Lx^2 + Ny^2 + z[-2 + \tfrac{4}{3}\left(\frac{4-3k}{1-k}\right)\mathfrak{C}'x + \tfrac{4}{3}\left(\frac{4-3k}{1-k}\right)\mathfrak{B}'y + k_4 z] = 0 \tag{41}$$

k_4 being a parameter.[18]

---

[18] By means of the Fubini canonical differential equations of a surface S referred to its asymptotic net we can prove that the pencil (41) of quadrics associated with the tangents of the curves of a conjugate net at an ordinary point P_x is in the family of conjugal quadrics of S at P_x (V. G. Grove, "The transformation of Čech," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 231-234). Furthermore, if we define the two cross ratios in equations (39) respectively by two different constants k_1 and k_2 instead of k, we may obtain a family of quadrics, which is an extension of the pencil (41) but belongs to the more general family defined by Grove ("Quadrics associated with a curve on a surface," *Bulletin of the American Mathematical Society*, vol. 51 (1945), pp. 281-287).

Observing that any quadric having contact of the second order with the surface S at the point P_x is given by

$$Lx^2 + Ny^2 + z[-2 + k_2x + k_3y + k_4z] = 0, \tag{42}$$

k_2, k_3, k_4 being arbitrary parameters, we infer immediately that the quadrics of the pencil (41) belong to this family.

If $k = \frac{1}{3}$ and Q_k passes through the point P_y, we have the canonical quadric of Davis [1, p. 10]; if $k = \infty$, we have the quadrics of Darboux; and if $k = 0$, we have the pencil [8, p. 697] of the quadrics having contact of the third order with the parametric curves at the point P_x of the surface S. Thus we have proved the

THEOREM. *The correspondence C_0 is the polarity with respect to the pencil of the quadrics having contact of the third order with the parametric curves at the point P_x of the surface S.*

Since the polar line of the axis with respect to the quadric Q_k is a general canonical line of the second kind of Davis,

$$z = \tfrac{2}{3}\left(\frac{4-3k}{1-k}\right)\mathfrak{C}'x + \tfrac{2}{3}\left(\frac{4-3k}{1-k}\right)\mathfrak{B}'y - 1 = 0, \tag{43}$$

and the polar line of the ray with respect to the quadric Q_k is a general canonical line of the first kind of Davis,

$$Lx + \tfrac{2}{3}\left(\frac{4-3k}{1-k}\right)\mathfrak{C}'z = 0, \qquad Ny + \tfrac{2}{3}\left(\frac{4-3k}{1-k}\right)\mathfrak{B}'z = 0, \tag{44}$$

we obtain the [2]

THEOREM. *The polar lines of the axis and the ray with respect to the quadric Q_k are respectively the canonical lines* (43) *and* (44) *of the second and the first kinds of Davis which depend upon the quadric Q_k.*

By a proper selection of the constant k these lines may be made to become any desired canonical lines of both kinds of Davis. In particular, when $k = \infty$, the lines (43) and (44) are respectively the associate ray and the associate axis. When $k = 0$, we obtain the principal join [3] and the canonical line $l_1(\frac{2}{3})$ of the first kind.[4]

---

[2] A general canonical line of each kind of the surface referred to the asymptotic curves has been characterized geometrically in a similar manner. See my paper [3].

[3] For the definition of the principal join see E. P. Lane [5, p. 702].

[4] These two characterizations were given by Lane and MacQueen [8, pp. 698-699].

Moreover, the polar line of any canonical line $l_1(k_1)$ of the first kind with respect to the quadric Q_k is a canonical line $l_2(k_2)$ of the second kind which depends on the selection of the constant k; and the values of k_1 and k_2 are the same when and only when $k = \frac{4}{3}$, that is, when and only when Q_k is a quadric in the canonical pencil of Davis.[5]

6. Two one-parameter families of lines. The equation of the family (41) of quadrics may be rewritten in homogeneous coordinates as

$$(45) \qquad Lx_2^2 + Nx_3^2 + x_4(-2x_1 + k\mathfrak{C}'x_2 + k\mathfrak{B}'x_3 + k_4x_4) = 0,$$

where k is a new constant. The polar line of the line $l_1(33)$ with respect to a general quadric of the family (45) is found to intersect the u-, v-tangents through the point P_x respectively in the points

$$(46) \qquad \Phi_k = (k\mathfrak{C}' - 2eL)x + 2\rho, \qquad \Psi_k = (k\mathfrak{B}' - 2\bar{e}N)x + 2\sigma.$$

On the other hand, the homogeneous coordinates of a point X near P_x and on the u-curve through P_x can be expressed as power series in the increment Δu corresponding to displacement from P_x to the point X along the u-curve, namely [8, p. 693],

$$(47) \qquad \begin{aligned} x_1 &= 1 + b\Delta u + \cdots, \\ x_2 &= \Delta u + \tfrac{1}{2}a\Delta u^2 + \cdots, \\ x_3 &= \tfrac{1}{6}(\mathfrak{H}/r)\Delta u^3 + \cdots, \\ x_4 &= \tfrac{1}{2}L\Delta u^2 + \tfrac{1}{6}L(a + b - l_u)\Delta u^3 + \cdots. \end{aligned}$$

For the purpose of finding the projection of the u-curve onto its osculating plane from any point τ on the line l_1:

$$(48) \qquad \tau = \lambda x - e\rho - \bar{e}\sigma + y,$$

λ being an arbitrary scalar function of u, v, and in connection with the later development, it is convenient to introduce another local tetrahedron with vertices at the points x, Φ_k, τ, y. If the point in space with coordinates $x_1, \cdots, x_4$ referred to the tetrahedron x, ρ, σ, y has coordinates $X_1, \cdots, X_4$ referred to the new tetrahedron x, Φ_k, τ, y, then the identity

$$x_1x + x_2\rho + x_3\sigma + x_4y = X_1x + X_2\Phi_k + X_3\tau + X_4y$$

[5] By means of this property Rasmussen and Hagen have defined the quadrics of the pencil (41), but they have made an error in considering k_1 and k_2 as always the same. See their joint paper [9, p. 301].

yields the equations for *the transformation of coordinates* between the two tetrahedra; after solution for $X_1, \cdots, X_4$ these equations can be written in the form

$$
\begin{aligned}
&X_1 = x_1 + (eL - \tfrac{1}{2}k\mathfrak{C}')x_2 + 1/\bar{e}[e(\tfrac{1}{2}k\mathfrak{C}' - eL) + \lambda]x_3,\\
&X_2 = \tfrac{1}{2}x_2 - (e/2\bar{e})x_3, \qquad X_3 = -(1/\bar{e})x_3,\\
&X_4 = (1/e)x_3 + x_4.
\end{aligned} \tag{49}
$$

The parametric equations of the projection C'_u of the u-curve from the new vertex $(0, 0, 1, 0)$ onto the tangent plane, $X_3 = 0$, are found by substituting the series (47) for $x_1, \cdots, x_4$ into equations (49) and taking such a linear combination of the resulting coordinates $X_1, \cdots, X_4$ and of $0, 0, 1, 0$ as will make the third coordinate vanish. These parametric equations, to terms of as high degree as will be needed, are

$$
\begin{aligned}
&X_1 = 1 + (b + eL - \tfrac{1}{2}k\mathfrak{C}')\Delta u + \cdots,\\
&X_2 = \tfrac{1}{2}\Delta u + \tfrac{1}{4}\alpha\Delta u^2 + \cdots,\\
&X_4 = \tfrac{1}{2}L\Delta u^2 + \tfrac{1}{6}L(\alpha + b - l_u + \mathfrak{H}/\bar{e}N)\Delta u^3 + \cdots,
\end{aligned} \tag{50}
$$

the third coordinate being zero in the remainder of this section. Imposing on the general equation of a conic the conditions that it be satisfied by the series (50) for X_1, X_2, X_4 identically in Δu as far as the terms in Δu^3, we obtain *the equation of the four-point conics at the point P_x of the projection C'_u of the u-curve,* namely,

$$
X_1X_4 - 2LX_2^2 - 2[eL + \mathfrak{H}/3\bar{e}N + (16 - 3k)\mathfrak{C}'/6]X_2X_4 + h_4X_4^2 = 0, \tag{51}
$$

where h_4 is a parameter. The polar line of the point $\Phi_k(0, 1, 0)$ with respect to any one of these conics has the equation

$$
2LX_2 + [eL + \mathfrak{H}/3\bar{e}N + (16 - 3k)\mathfrak{C}'/6]X_4 = 0. \tag{52}
$$

Since the plane of the line l_1(33) and the v-tangent intersects the osculating plane of the u-curve at the point P_x in the line

$$
X_3 = 0, \qquad 2X_2 + eX_4 = 0, \tag{53}
$$

we know that the line (52) coincides with the line (53) in case

$$
\bar{e} = 2\mathfrak{H}/(3k - 16)N\mathfrak{C}'. \tag{54}
$$

Similarly, with the roles of the curves u, v interchanged, we get

$$
e = 2\mathfrak{K}/(3k - 16)L\mathfrak{B}'. \tag{55}
$$

Thus for each pencil of quadrics of the family (45) we can determine two lines, namely, the line $l_1(33)$ satisfying the conditions (54), (55) and its polar line with respect to the pencil of quadrics, and hence *associated with the point P_x of the parametric conjugate net N_x there exist two one-parameter families of lines,* which will be called, respectively, *the first and the second families.* In particular, a reference to the original definition of Green [2, p. 114] shows that when $k = 4$ the two lines of these families may be regarded as generalizations of the canonical edges of Green of the asymptotic net of a surface to a conjugate net. Further, we observe that *all lines of the first family lie in a plane,* whose equation is found to be

$$\mathfrak{H}\mathfrak{B}'x_2 - r\mathfrak{K}\mathfrak{C}'x_3 = 0. \tag{56}$$

Finally, noticing equations (31), (32) it is easy to conclude that *if one of the two families of lines mentioned above is a canonical pencil of Davis, then the other is also, and necessary and sufficient conditions for this are*

$$\mathfrak{H} = \mathfrak{K} = m\mathfrak{B}'\mathfrak{C}' \qquad (m = \text{const.} \neq 0). \tag{57}$$

MICHIGAN STATE COLLEGE AND
NATIONAL UNIVERSITY OF CHEKIANG.

BIBLIOGRAPHY.

1. W. M. Davis, "Contributions to the theory of conjugate nets," *Chicago doctoral dissertation* (1932).
2. G. M. Green, "Memoir on the general theory of surfaces and rectilinear congruences," *Transactions of the American Mathematical Society*, vol. 20 (1919), pp. 79-153.
3. C. C. Hsiung, "The canonical lines," *Duke Mathematical Journal*, vol. 8 (1941), pp. 738-742.
4. C. C. Hsiung, "The contact of conics and quadrics with two hypersurfaces in space of n dimensions," to appear elsewhere.
5. E. P. Lane, "Invariants of intersection of two curves on a surface," *American Journal of Mathematics*, vol. 54 (1932), pp. 699-706.
6. E. P. Lane, "A canonical power series expansion for a surface," *Transactions of the American Mathematical Society*, vol. 37 (1935), pp. 463-482.
7. E. P. Lane, *A Treatise on Projective Differential Geometry*, The University of Chicago Press, 1942.
8. E. P. Lane and M. L. MacQueen, "The curves of a conjugate net," *Duke Mathematical Journal*, vol. 5 (1939), pp. 692-704.
9. R. B. Rasmussen and B. L. Hagen, "Comments on canonical lines," *Bulletin of the American Mathematical Society*, vol. 47 (1941), pp. 298-302.

SIX-RINGS IN MINIMAL FIVE-COLOR MAPS.*

ARTHUR BERNHART.

Introduction. Kempe [1] and Heawood [2] have shown that five colors are sufficient for coloring any map on a sphere. We conjecture the existence of some maps for which five colors are necessary, and seek the simplest example. For this end *we here consider only minimal maps,* namely those five-color maps such that any map with fewer regions is four-colorable. This paper investigates the structure of minimal maps by a systematic analysis of rings.

By a *proper n-ring* we mean [1] a cycle of n distinct regions, each adjacent to the regions which precede and succeed it in the cyclic order, [2] but to no other region in the cycle, and [3] dividing the rest of the map into two non-empty sides. This definition is equivalent [in minimal maps] to the usage of Birkhoff [3] who introduced the terminology " ring of n regions." The first condition conveys the generic meaning of the word ring as used by many four-color writers without a formal definition. The second condition originated with Birkhoff, and corresponds to what Veblen [4] intended by a *simple* circuit. For $n = 1, 2, 3$ it adds nothing to the generic meaning, but Birkhoff excluded these cases thereby implying the third condition. A ring divides the regions of a map into three mutually exclusive parts: the regions R of the ring itself, the regions I inside the ring, and the regions O outside the ring. Ordinarily the terms " inside " and " outside " are interchangeable, but whenever the structure of one side is simpler or more fully known we shall refer to it as the inside, achieving thereby an economy of description. In proper rings, I is a proper subset of the map. Whenever we use the term ring in the generic sense with a meaning other than that set forth in the foregoing definition, we shall warn the reader by calling it an *improper* ring. Thus a single region forms an improper 1-ring, two adjacent regions form an improper 2-ring, the regions meeting at a vertex form an improper 3-ring, and the regions participating in an edge form an improper 4-ring. Conversely, these examples of improper n-rings are the only possibilities for

* Received August 8, 1946.

[1] A. B. Kempe, *American Journal of Mathematics*, vol. 2 (1879), pp. 193-200.

[2] P. J. Heawood, *Quarterly Journal of Mathematics*, vol. 24 (1890), pp. 332-338.

[3] G. D. Birkhoff, *American Journal of Mathematics*, vol. 35 (1913), pp. 115-128.

[4] O. Veblen, *Annals of Mathematics*, vol. 14 (1912-13), pp. 86-94.

$n = 1, 2, 3, 4$ respectively. In this paper unless the context indicates otherwise, all maps are minimal, and all rings are proper.

The most significant result is the 6-ring theorem which lists the colorability of each side of a 6-ring giving all solutions compatible with the Kempe chain reductions. Prior to the main theorem the case of n-rings with $n < 5$ is solved again, but in addition to the familiar result that such rings are necessarily improper, a new conjugate-edge theorem is demonstrated. Birkhoff's unique solution for 5-rings is obtained algebraically by solving a simultaneous set of primary and contingent options. This proof serves as a model for subsequent theorems, whose proofs are more involved but embody the same argumentative procedure. A compact notation is developed such that one "color matrix" of sixteen elements expresses 96 contingent options of the Kempe type.

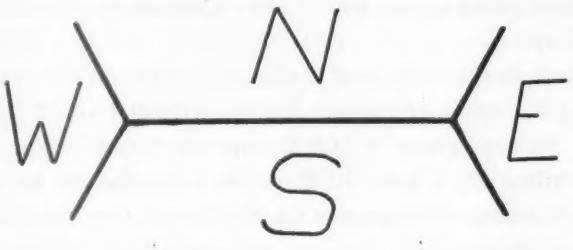
In order to make the discussion of 6-rings more readable, the step-by-step proof of the main theorem has been relegated to a separate section at the end. The significance of the main theorem is discussed. Whereas four-color investigation usually begins with geometric structures and derives their colorability, the methods used here permit arguing from colorability to structure. This innovation proves fruitful, and the main theorem is used to obtain answers to seven structural questions. Each of these results may be read as a theorem in itself revealing new reducible configurations. The Birkhoff result that "a ring of six surrounding four regions is reducible" is extended to a 6-ring surrounding n regions, $3 < n < n_0$ where n_0 is at least 16. Finally, if a 6-ring surrounds j inside regions ($j > 3$), composed of p peripheral regions making contact with the ring plus h additional regions ($j = p + h$), then there are less than $j - 5$ inside pentagons, the p regions form a p-ring with $p > 8$, and $\Sigma(n - 5) = h > 6$, where the summation extends over the j inside n-gons.

4-rings.

Of the four regions around an edge the two which are adjacent will be designated as the *primary components,* the other two the secondary or *guard* components. Let regions N and S be adjacent guarded by regions E and W. If the inside structure of this 4-ring is rearranged to make E and W adjacent thereby separating N and S, the primary and secondary components of the edge are interchanged, and the new edge is *conjugate* to the old.

As a non-trivial corollary to the study of improper 4-rings we state the result for any minimal map: *If one edge is replaced by its conjugate, the*

new map can be 4-colored. This result shows that each map is on the *verge*



of colorability and the slight modification necessary to permit 4-coloring can be achieved *without reducing the number of regions.*

Though the proof is simple we include it here, since the procedure furnishes a prototype for the cases $n > 4$. The entire map cannot be 4-colored, but the outside regions may or may not be colorable consistent with a specified *color scheme* for the ring. We use 1 to indicate one color, 2 a color different from 1, 3 a color different from 1 and 2, 4 a color different from 1, 2 and 3, listing the colors for *SENW* respectively. A 4-ring can be colored in only four schemes, scheme A: 1212, scheme B: 1213, scheme C: 1232, scheme D: 1234. On coloring R according to any scheme S we use the superscript shorthand S^1 inside colorable, S^2 outside colorable, S^3 outside not colorable, S^4 inside not colorable, S^5 neither side colorable, and S° both sides colorable. We use a *dot* or mere *juxtaposition* for the logical product (both-and) of two propositions, and a *comma* for the logical option (and/or). When an argument is equally valid for inside and outside the superscripts may be omitted. Thus (S^1, S^4) and (S^2, S^3) are logical options required by the law of excluded middle, while S^1S^4 and S^2S^3 are contradictions; $S^3S^4 = S^5$ and $S^1S^2 = S^\circ$ but S° is absurd since by definition minimal maps are not four-colorable.

Since the primary components N and S are adjacent, the edge itself can be colored only by schemes C and D. In the shorthand notation $(CD)^1(AB)^4$. Now if C^2, then $C^1C^2 = C^\circ$, which is absurd. Therefore C^3. Likewise D^3. But deleting the edge between N and S without disturbing the outside structure forms a map with fewer than the minimal number of regions, and hence this modified map can be 4-colored, therefore (A^2, B^2). We refer to this as a primary or *reducing option.* But Kempe [1] has pioneered the way in showing that if the side of an n-ring is colorable according to one color scheme (on the bounding ring) then it must be colorable also according to certain other

schemes. Thus $A \rightarrow (B, C)$; $B \rightarrow (A, D)$; $C \rightarrow (A, D)$; and $D \rightarrow (B, C)$. We refer to such implications as secondary or *contingent options*. The option (B, C) may be satisfied by any one of three alternatives: B and not-C, C and not-B, both B and C.

Kempe derived the contingent option $D : 1234 \rightarrow (B : 1214, C : 1232)$ by introducing the notion of chains. Regions colored 1 and 3 form an "odd" league, while regions colored 2 and 4 form an "even" league. If—under scheme D—regions $S : 1$ and $N : 3$ can be connected by an uninterrupted chain of outside regions belonging to the odd league, then this chain partitions the outside and its bounding ring into two subdivisions. (The chain regions may be assigned arbitrarily to either subdivision.) In one subdivision Kempe retained the original colors, but in the other subdivision he interchanged colors 2 and 4. This exhibits a new 4-coloring of the outside subject to a correspondingly new color scheme on the ring, in this instance scheme C.

But if no such odd-chain connecting S and N exists, Kempe argued that E and W must be connected by an even chain. Interchanging 1 and 3 in one partition formed by this chain, he obtained scheme B, for 1214 and 1213 are dual notations for the same scheme.

Besides the leagues 1-3 versus 2-4 we could use 1-2 vs 3-4 or 1-4 vs. 2-3. The same topological argument is applicable and can be extended to rings $n > 4$, with precisely similar results for schemes with odd-even-odd-even ring coloring. If the ring coloring degenerates to a single odd-even pattern, as in $B : 1213$, the argument is sterile, but in the general case three contingent options are obtained by chain deduction from each postulated scheme.

In the improper 4-ring case under consideration, the four contingent options can be diagrammed by putting the schemes at the corners of a square, A and D placed diagonally. Then each scheme implies an adjacent corner. Since C^3D^3, from the option (A^2, B^2) we obtain $(A^2 . B^2)$ and in particular B^2. This establishes our conjugate-edge theorem.

If the 4-ring were proper we would have four primary options $(A, B)(A, C)(C, D)(B, D)$ for both sides, which can be satisfied only by putting A, D and $B . C$ on opposite sides. But this is contrary to the contingent options, hence proper 4-rings are excluded from minimal maps.

The exclusion of proper 1-rings, 2-rings, and 3-rings is simpler since it does not require chain deductions. For in each case only one color scheme is possible for the ring. Primary options compel the coloring of each side by at least one scheme, and with both sides colorable with the same scheme, the map would be 4-colorable.

5-rings.

Let regions A, B, C, D, E constitute a proper 5-ring. Only ten schemes are possible: scheme A with $ABCDE$: 12323 respectively; scheme A^* with $ABCDE$: 12342; scheme B with $BCDEA$: 12323; etc. The notation emphasizes the cyclic symmetry. Scheme A uses only three colors in which the color of region A is not repeated. Scheme A^* uses four colors in which region A lies between two regions colored the same.

We have the chain deduction A : 12323 $\rightarrow$ (E : 12123, D^* : 12343), and by clock symmetry $A \rightarrow (B, C^*)$. Again A^* : 12342 $\rightarrow$ (D : 12142, C^* : 12324) and by clock symmetry $A^* \rightarrow (C, D^*)$. Cyclic symmetry gives us twenty contingent options, two for each scheme.

Merging the inside with regions B and D we form a modified map with fewer than the minimal number of regions, and which therefore can be 4-colored. This yields the option (A^*, C, D). Merging region A with the inside yields (A, A^*, C^*, D^*). Symmetry provides a total of ten primary options, applicable to both outside and inside.

We seek all solutions for this system of 10 primary and 20 secondary options. Trying $(C.D)^2$ we are led directly to the solution $(A.B.C.D.E)^2$ $(A^*.B^*.C^*.D^*.E^*)^1$. In trying for other solutions we must avoid putting two "consecutive unstarred" schemes on the same side. But then the try $(A)^2$ implies $(A.C^*.D^*)^2(A^*)^1$ hence $(A^*.C.D)^1$ contrary to the primary option $(A^*, C, D)^2$. Avoiding $(A)^2$ and by symmetry also $(B)^2$, $(A)^1$ etc., we would have to exclude A, B, C, D, E from both sides. But then the option (A^*, C, D) would compel A^* for both sides, contrary to the definition of minimal maps. This shows that the given solution is the only one. It corresponds to an inside composed of a single pentagon.

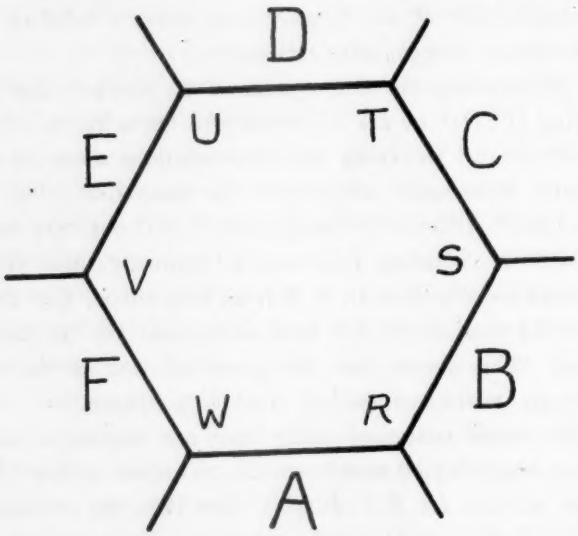
For if the inside contained more than one region, a merging of the inside to form a single region would provide the outside option $(A, B, C, D, E)^2$ leading to the solution $(A.B.C.D.E)^2$. But then the corresponding inside option $(A, B, C, D, E)^1$ could not be satisfied. Thus the case of an inside composed of more than one region is excluded. Hence the inside of a 5-ring must be a single region, a pentagon, and the outside must be colorable by each of the schemes using 4 colors on the ring, and by none other. It is to be noted that minimal maps must include 5-rings, one for each pentagon. By contrast the existence of 6-rings is only conjectural.

6-rings.

Let regions A, B, C, D, E, F constitute a proper 6-ring. Thirty-one schemes are possible:

$121212 = X$	$123123 = 0$	$123242 = H$
$213 = F$	$124 = 9$	$243 = C^*$
$232 = E$	$132 = 1$	$412 = T$
$234 = V$	$134 = D^*$	$413 = F^*$
$312 = D$	$142 = A^*$	$414 = S$
$313 = B$	$143 = 8$	$423 = 7$
$314 = G$	$212 = C$	$424 = E^*$
$323 = 2$	$213 = 3$	$432 = 4$
$324 = B^*$	$214 = 6$	$434 = R$
$342 = U$	$232 = A$	
$343 = 5$	$234 = W$	

The notation emphasizes the following cyclic groups $(ABCDEF)$ $(A^*B^*C^*D^*E^*F^*)(RSTUVW)(GH)(X)(0)(123)(456)(789)$.



We now derive 38 primary options, each applicable to both sides. Merging region A with inside: $(A, C^*, E^*, H, R, W, 4, 7)$, six variants. Merging regions A and D with inside: $(A^*, D^*, 0, 1, 8, 9)$, three variants. Merging regions B, D, F and inside: (A, C, E, H, X), two variants. Merging regions C, E and inside: $(A, D^*, R, W, 1, 4)$, six variants. Merging B and F, also C and E through inside: $(A, D, X, 1, 4)$, three variants. Merging B and F, also making C and E adjacent, through inside: (A^*, C, E, H, T, U), six variants. Partitioning inside into two regions by an edge guarded by A

and D, then merging these subdivisions of the inside with B and E respectively: $(A^*, D^*, R, U, 1,4, 5, 8)$, three variants, and clock symmetric: $(A^*, D^*, T, W, 1, 4, 6, 9)$, three variants. Finally, partitioning inside into three pentagons by edges guarded by B, D, F then annexing two pentagons to C and E, respectively: $(B^*, C, E, F^*, T, U, 0, 2, 3, 7)$, six variants. The modified map has fewer regions except when the inside consists of a single hexagon. For this exceptional case the option is justified for the outside by conjugating the edge guarded by C and E.

Merging the inside into a single hexagon yields $(X,A,B,C,D,E,F,0,1,2,3)$, an option not applicable to the outside of a $6\,R\,1$. Making two pentagons from the inside yields $(A^*, B, C, D^*, E, F, G, H, X, 5, 6, 8, 9)$, three variants, not applicable to the outside of $6\,R\,1$ or $6\,R\,2$. Making three pentagons with a common vertex from the inside yields $(A,B^*,C,D^*,E,F^*,G,H,R,S,T,U,V,W)$, two variants, not applicable to the outside of $6\,R\,1$, $6\,R\,2$, and $6\,R\,3$.

Next we list all chain deductions from the first scheme in each cyclic group.

$A \to (H, 1)$ and $\begin{pmatrix} C & E & X \\ R & W & 4 \end{pmatrix}$.

$A^* \to (H, 1)$ and $(U, 9)$ and $(T, 8)$.

$G \to (B, B^*)$ and (D, D^*) and (F, F^*).

$R \to (D^*, E^*)$ and $\begin{pmatrix} A & W & 4 \\ S & B & 5 \end{pmatrix}$.

$X \to \begin{pmatrix} A & C & E \\ D & F & B \end{pmatrix}$:

$0 \to (1, 7)$ and $(2, 8)$ and $(3, 9)$.

$1 \to (A, A^*)$ and (D, D^*) and $(0, 4)$.

$4 \to (1, 7)$ and $\begin{pmatrix} R & A & W \\ U & D & T \end{pmatrix}$:

$7 \to (B^*, C^*)$ and (E^*, F^*) and $(0, 4)$.

The ring pattern odd-even-odd-even-odd-even occurs for the first time in 6-rings and it involves more than two alternatives. Thus X : 121212 requires a selection of at least one scheme from each of six three-way choices (A, B, C) (B, C, D) (C, D, E) (D, E, F) (E, F, A) (F, A, B). We indicate all six of these by the matrix $\begin{pmatrix} A & C & E \\ D & F & B \end{pmatrix}$. These requirements can be satisfied in any one of five sufficient ways (AD) (BE) (CF) (ACE) and (BDF), namely *by any row or column from the matrix.*

The entire system of contingent options can be presented in a more compact manner. We may arrange the schemes which involve simple odd-even-odd-even ring patterns in groups of four,

AH	BG	CH	DG	EH	FG	10	20	30
$1A^*$	$2B^*$	$3C^*$	$1D^*$	$2E^*$	$3F^*$	47	58	69
B^*V	C^*W	D^*R	E^*S	F^*T	A^*U			
$7C^*$	$8D^*$	$9E^*$	$7F^*$	$8A^*$	$9B^*$			

in which each scheme implies another scheme at an adjacent corner. Those schemes which involve a threefold odd-even sequence around the ring we arrange in an array of 4 rows and columns corresponding to the design:

121212	3232	3212	1232	$X\,A\,C\,E$
3432	1412	1432	3412	$4\,D\,U\,T$
3214	1234	1214	3234	$6\,V\,F\,W$
1434	3414	3434	1414	$5\,S\,R\,B$

Each element in this square array is to be thought of as belonging to one row and to one diagonal. To obtain the 2 by 3 option matrix for any scheme we write the elements in the same diagonal below the elements in the same row, retaining the column pairings. Or we may read the contingencies directly from this square array. Thus scheme R implies one of five sufficient alternatives: all the schemes $(5\,S\,B)$ in the same R-row; all the schemes $(4\,A\,W)$ in the same R-diagonal; or one representative from the R-row and one from the R-diagonal chosen from the same column $(4\,5)$ $(A\,S)$ or $(W\,B)$.

The next task is to find all solutions for the system of primary and contingent options. From any one solution we can, of course, produce others by cyclic permutation, by clock symmetry, and by interchanging inside and outside. We shall count all these variations as belonging to the same solution. We are now in a position to state the theorem: *The 6-ring options admit only these six solutions*:

Solution #1 (1 inside variation, 10 outside variations)

Inside *ABCDEFX*0123

Outside *A*B*C*D*E*F*GH*456789 and two pairs $(RU)(SV)(TW)$ or more.

Solution #2 (3 variations, each side)

Inside *A*BCD*EFGHX*5689

Outside *AB*C*DE*F*RTUW*012347

Neither side *SV* .

Solution #3 (2 variations, each side)

Inside *AB*CD*EF*GHRSTUVW*

Outside *A*BC*DE*FX*0123789

Neither side 456 .

Solution #4 (3 variations, each side)
Inside $AB^*C^*DE^*F^*GHSVX456789$
Outside $A^*BCD^*EFRTUW0123$.

Solution #5 (6 variations, each side)
Inside $AB^*CD^*EF^*GHRTUW47$
Outside $A^*BC^*DE^*FSVX01235689$.

Solution #6 (12 variations, each side)
Inside $ABC^*DE^*F^*GHRSX4578$
Outside $A^*B^*CD^*EFTUVW012369$.

Comments on these solutions: Solution #1 corresponds to a single hexagon inside; solution #2 corresponds to two pentagons inside, whose common edge in variation one is guarded by A and D; solution #3 corresponds to three pentagons inside, with a common inside vertex, whose edges in variation one are guarded by A, C, E. That is, we know the inside structure of the first three solutions and it is a simple matter to verify that such structures are 4-colorable only for the schemes listed. We do not know the outside structure, and it is a matter of conjecture whether such structures, colorable as indicated, actually exist.

An attempt has been made to list the solutions in the order of their simplicity, but since the geometric structure of the last three solutions is unknown, the assigned order is merely for reference. Likewise the distinction between inside and outside is arbitrary. Solution #5 (with schemes 5 and 6 inadvertently omitted) was given by Birkhoff.[3] He was disappointed in finding more than three solutions. But now that we know all the color solutions for the 6-ring, we can settle many questions without knowing their inside geometrical structure. In examining one side of a 6-ring we may consider in turn each of the 63 possible variations. Symmetry or other considerations may reduce the number of cases requiring separate study. The important contribution is that we have a tool for investigation—a partial "analytic geometry" for maps—which frees us from the necessity of drawing diagrams. We use geometric intuition to set up the options, but their simultaneous solution is a problem in permutations and combinations. Furthermore, our solutions are *complete* in the sense that each satisfies all the conditions proposed; and so no further results can be obtained from geometric devices which do not arrive at new conditions beyond those contained in our options. Our solutions are incomplete in that our options may not express all the conditions inherent in the geometric formulation of the problem.

Since the 6-ring theorem lists *all* solutions, we may set down a number of simple corollaries which summarize their characteristics. Observe that A and A^* are always on opposite sides; and that $(G.H)$ are always opposite $(0.1.2.3)$. Also notice the absentees 4. 5. 6 in #3, and SV in #2, optionally absent in #1. The only solution clock-asymmetric is #6. The only solution with three-phase symmetry is #3.

Inside Structure for 6-rings—Preliminary Considerations.

Let $6Rn$ indicate that a ring of six regions surrounds an inside group of n regions. It has already been mentioned that the first three solutions of the 6-ring correspond to $6R1$, $6R2$ and $6R3$ respectively.

If both $6Rn'$ and $6Rn''$ have the same 6-ring solution, then $n' = n''$. For if n' is not greater than n'' then the inside with n' regions could be matched with the outside with $m_0 - n'' - 6$ regions forming a non-colorable map with $m = m_0 - (n'' - n')$ regions. But m_0 is the minimal value of m, therefore $n' = n''$. Except for the ambiguous $6R1$ outside, if a geometric structure can be found fitting one variation, it could be rotated and reflected to fit all the other variations. It is conceivable that more than one geometric structure (different arrangements of the same number of regions) might fit the same solution, or that no such structure was geometrically possible. Again, such structures might be possible, yet not essential, features of every minimal map.

Knowing the color solutions of the six possible 6-rings, it is natural to seek geometric structures meeting these requirements. But the trial and error method of examining one geometric candidate after another seems to be unfruitful. Birkhoff [3] cautiously confessed: "In all the cases which I have considered, the ring R of six regions containing more than three [inside] regions . . . is reducible." (The terminology *reducible* means that the problem of coloring such a map reduces to the problem of coloring another map with fewer regions. In minimal maps reducible configurations simply do not occur.) We propose an approach to 6-ring structure by means of a cumulative series of questions. Excluding from consideration the $6R1$ inside, $6R2$ inside, and $6R3$ inside, whose structures are already completely known, our questions will apply to each side of all six rings. Each such side must be colorable according to one of the variations listed as possible in the foregoing 6-ring theorem. Since the color schemes of each variation are known, we can decide whether its colorability is compatible with the conditions in question. When a like verdict is obtained for each of the variations, we have obtained a theorem valid for all 6-rings. If the verdict should depend upon

the variation, then we have a clue to the characteristics which distinguish one variation from another. Each successive answer will suggest the formulation of new questions. This is our program. The questions discussed in this paper are by no means exhaustive, but they will serve to illustrate the fruitfulness of this approach.

Can any inside region Y make multiple contact with the 6-ring? To avoid taboo 2, 3, 4-rings contact must be at opposite sides, say A and D. Then $ABCDY$ and $AYDEF$ are 5-rings, so that their insides X and Z are single regions. But then $AXDZ$ forms a taboo 4-ring around Y. Hence *multiple contact is excluded,* and the peripheral cycle of inside regions making contact with the ring is a *distinct* set.

Can any inside region contact as many as three ring regions? Since multiple contact has been excluded, the contact must be consecutive. If D' contacts $ABCDE$ then $AD'EF$ threatens to form a 4-ring unless D' is the only inside region, hence $6\,R\,1$. If D' contacts only $ABCD$ then $AD'DEF$ forms a 5-ring around D'', and $D'D''$ are the two inside pentagons for $6\,R\,2$. If D' contacts only CDE then both $ABCDEF$ and $ABCD'EF$ would be 6-rings with DD' forming a *double link.*

Designating the respective insides by I and J we observe that I contains one more region D' than J. But the colorability of I and J are interdependent. Postulating scheme C for one side implies T for the other side, and conversely. The same reciprocity holds for E and U, A^* and H, B^* and 2, C^* and 8, E^* and 9, F^* and 3, S and 6, V and 5, 0 and 7. Of the three schemes B, F, G postulating one for I implies an option of the other two for J, while postulating one for J implies both the other two for I. (Consequently postulating one of the three schemes for I implies a second for I, and the third for J.) The same implications hold for the triads A, 1, 4 and D^*, R, W. Finally postulating X for one side implies D for the other, but $I(D) \rightarrow J(D, X)$ while $J(D) \rightarrow I(D.\,X)$, relations which could be obtained formally by treating D, D, X as a fourth triad.

In view of these interdependent relations between I and J we may answer the double-link question by systematically trying all 6-ring solutions seeking an IJ fit.

$J(0, 1, 2, 3) \rightarrow J$ (0. 1. 2. 3. not-G. not-H) $\rightarrow I\,(AB^*F^*47) \rightarrow J\,(B^*F^*)$ since B^* is equivalent to not-B, F^* to not-F. Thus we can ignore any postulate $J(0, 1, 2, 3)$ not coupled with $J(B^*F^*)$. The alternate postulate $J(GH) \rightarrow I(BFA^*)$ whence J is not-4. This excludes many J cases, in particular both sides of $6\,R\,1$ and both sides of #4. Since $C.\,C^*$ never occur simul-

taneously, we avoid putting T and 8 on the same side, J or I; likewise $E.E^*$ indicates U and 9 are incompatible; while $D.D^*$ indicates X or D should not be placed on J with either R or W. Only six cases for J pass these preliminary hurdles, and we examine these individually.

First case: J is #2 inside, first variation, with $A^*BCD^*EFGHX5689$. Then I is $A^*BC^*DE^*FGHRSTUVW$, which is #3 inside, second variation.

Second case: J is #3 inside, first variation, with $AB^*CD^*EF^*GHRSTUVW$. Then I is $A^*BCD^*EFRTUW1456$ which does not fit any 6-ring solution.

Third case: J is #3 outside, second variation, with $AB^*CD^*EF^*X0123789$. Then I is $AB^*C^*DE^*F^*TU0147$ which fits #2 outside, first variation. This is the first case as seen from the other side.

Fourth case: J is #5 inside, third variation, with $AB^*CD^*EF^*GHTVWS69$. Then I is $A^*BCD^*E^*FRSTUW123456$, which does not fit any 6-ring solution.

Fifth case: J is #5 inside, fifth variation, with $AB^*CD^*EF^*GHRSUV58$. Then I is $A^*BC^*D^*EFRTUVW123456$ which does not fit.

Sixth case: J is #5 outside, fourth variation, $AB^*CD^*EF^*SVX01235689$. Then I is $AB^*C^*DE^*F^*RSTUVW1234567$, which does not fit.

Summary: In any 6 R 3 inside, each of the enclosed pentagons produces a double link, forming three associated 6 R 2 structures. Otherwise, *6-rings do not possess double links*; no region contacts more than two ring regions. The regions in the outermost I shell belong to two types: *corner regions* which contact two ring regions, guarding their common edge; and *lateral* regions which contact only one ring region. Incidentally, since this shell certainly contains six corners, 6 R 4 *and* 6 R 5 *structures are excluded.*

If there are p regions in this peripheral I shell (*p-shell*: composed of *corner* regions and *lateral* regions) and h additional regions on the inside, then the average number of sides for inside regions exceeds 5 by $\frac{h}{p+h}$, so that if they were only pentagons and hexagons there would be p pentagons and h hexagons.

Can non-consecutive regions in the p-shell be adjacent? Having shown that the regions in the outermost shell are distinct, we show that they constitute a ring! The only contact between non-consecutive regions in the shell not forming taboo 3-rings, 4-rings, or 5-rings would be between lateral

regions at opposite sides of the hexagonal contour. Label the corner regions R at AB, S at BC, T at CD, U at DE, V at EF, W at FA. Let A' be in lateral contact with A, and D' with D. Postulating A' adjacent to D' forms a *system of three* 6*-rings*: $ABCDD'A'$ with RST inside, $AA'D'DEF$ with UVW inside, and $AFEDCB$ with $WA'R$ outside. Designate the respective insides by I, J, K. *Their colorability is interdependent.* Thus $I(A)$ and $J(C)$ together imply K(not-E) else the entire map would be 4-colorable. In the ordered triad ACE, postulating the first scheme for I and the second scheme for J implies the exclusion of the third scheme from K. Symmetry in the definition of I, J, K allows the corresponding inferences from the triads CEA and EAC, (but not from ECA). 001 forms another such triad, and since the schemes 0 and 1 are always opposite H, we conclude that at least one of the sides I, J, K is colorable H. We next try to assign color schemes to I, J, K consistent with the six known solutions but avoiding respective assignments which involve taboo triads such as $I(A).J(C).K(E)$. The method is simple but involves a consideration of many cases, and we omit the details here. The conclusion from the systematic trials is that no fit is possible: the postulate A' adjacent to D' is absurd. Therefore *the p-shell constitutes a p-ring.*

In an entirely similar manner we may dispose of the question: *Can opposite corner regions W and T have a common neighbor X?* Another system of 6-rings is formed: $DEFWXT$, $TXWABC$, and $CBAFED$. Another system of scheme triads is formed such as OAD and again we may try for a fit. No fit can be made and the postulated structure is impossible.

Since the six or more regions in the first shell form a ring each must be adjacent to some region in the second shell. Since the same region X cannot serve for both W and T there must be at least two regions in the second. (For $p = 6$ we may apply our 6-ring conclusions to the p-ring. The trials $pR1$, $pR2$, $pR3$ collapse, whence there would be a p'-ring within the p-ring.) That makes at least nine regions inside the original 6-ring, so we have *excluded all* $6\,R\,6$, $6\,R\,7$ *and* $6\,R\,8$ *structures.*

Can the p-shell be another 6*-ring,* contiguous with the original 6-ring and consisting of corner regions only? Let the inside I of $ABCDEF$ include R and let the inside J of $RSTUVW$ exclude A. Then the color schemes of I and J are interdependent. Thus $J(B^*) \rightarrow I(A^*B^*C^*D^*0189)$, but since no 6-ring solution contains 0 together with four "consecutive star" schemes, the postulate $J(B^*)$ is absurd, and by symmetry *all* star schemes are excluded from J. But then J would have to be the $6\,R\,1$ inside, colorable by scheme 0. But $J(0)$ leads to the impossible $I(H.0)$, so that *contiguous* 6*-rings do not occur.*

Can the p-shell be a 7-ring? We suppose a region A' making lateral contact with A, plus the six corner regions. Then $A'RSTUVW$ is a 7-ring contiguous with the 6-ring. Considering each of the 63 cases, the color schemes not possible for that side of the 6-ring including A' induce restraints on the coloring of the 7-ring. But too many restraints become inconsistent with the requirements of certain primary options for the 7-ring. In this manner we may dispose of each case. Their verdict is unanimous, so that *a 7-ring contiguous with a 6-ring does not occur.*

In like manner minimal maps reject the occurrence of an 8-ring contiguous with a 6-ring. The 8-ring possesses two *lateral* regions in addition to the six *corner* regions. We distinguish four structures: (1) A' and A'' sharing a vertex with A, (2) A' and B' in lateral contact with A and B, respectively, (3) A' and C' in lateral contact with A and C, respectively, and (4) A' and D' in lateral contact with A and D, respectively. An exhaustive examination of each case shows that no color fit is possible, and the proposed structure is excluded.

We conclude our current study of the inside structure of 6-rings with the question: *Can alternate corner regions S and U have a common neighbor X in the second shell?* If so, then $UXSCD$ forms a 5-ring around the pentagon T. The colorability of the 7-ring $ABSXUEF$ and of the 6-ring $ABCDEF$ are interdependent. Let I be that side of the 6-ring which excludes X, and let J be that side of the 7-ring which excludes T. To each scheme for which I can be colored there are certain schemes excluded from J, lest the entire map be 4-colorable, contrary to our basic conjecture. Usually each I scheme excludes four J schemes, but $B^*, E^*, 0, 2, 7, 9$ each exclude six J schemes, while $I(X)$ excludes only two J-schemes. An examination of all 63 cases for I shows that only one case is compatible with the supposed structure, solution #1 inside, that in which I is a single hexagon. Then C and D are pentagons, the primary components of an edge guarded by the pentagon T and the hexagon I.

Franklin [5] has shown that three pentagons and one hexagon cannot form an edge with the hexagon in primary position. But when the hexagon is in guard position, the $7R4$ structure is *irreducible* since its outside may be assigned color schemes compatible with all known primary and contingent options.

The reader is reminded that these cumulative results concerning the

[5] P. Franklin, *American Journal of Mathematics*, vol. 44 (1922), pp. 225-236.

"inside structure" of 6-rings apply to *both* sides of solutions #4, #5, #6 and—with specified exceptions—to the *outside* of solutions #1, #2, #3.

Neighboring the original six-ring R are two more rings each containing at least nine regions. The p regions of each associated ring consist of six *corner* regions (guarding the edges between consecutive R regions) plus three or more *lateral* regions. Inside the p-ring a corner n-gon contacts $n-4$ regions, while a lateral n-gon contacts $n-3$ regions. Opposite corner regions have no common neighbor, and alternate corner regions have no common neighbor inside the p-ring. Defining the power of an n-gon as $n-5$, the combined power of the $p+h$ regions inside the 6-ring is precisely h. By resorting to geometrical trials it can be shown that the power h cannot be concentrated in five or fewer regions, nor yet in six hexagons. This includes the (weaker) corollary that h is greater than six. Combined with the knowledge that p is at least nine, we see that $p+h$ exceeds fifteen. Thus *6Rn structures with* $3 < n < n_0 = 16$ *do not exist* in minimal maps. The way is open to increase the value of p, h and n_0. Birkhoff [3] conjectured that six-rings not surrounding one, two or three regions might be a characteristic feature of all minimal maps. If one such ring is present in one minimal map of m_0 regions then m_0 is not less than $2n_0 + 6$.

Proof of the Main Theorem.

The superscript notation relative to scheme S provides for five classes:

S^1 inside colorable $\quad S^3$ outside not-colorable $\therefore (S^1, S^5)$

S^2 outside colorable $\quad S^4$ inside not-colorable $\therefore (S^2, S^5)$

S^5 neither side colorable

$(RST)^2 = R^2 . S^2 . T^2$

The problem is solved when each of the 31 schemes has been assigned one of the superscripts 1, 2, or 5 consistent with the primary and contingent options.

For convenient reference we tabulate herewith all the *primary* options:

(1)	(A, C, E, H, X)	(21)	$(A, D^*, W, R, 1, 4)$
(2)	(B, D, F, G, X)	(22)	$(B, E^*, R, S, 2, 5)$
		(23)	$(C, F^*, S, T, 3, 6)$
(3)	$(A, D, X, 1, 4)$	(24)	$(D, A^*, T, U, 1, 4)$
(4)	$(B, E, X, 2, 5)$	(25)	$(E, B^*, U, V, 2, 5)$
(5)	$(C, F, X, 3, 6)$	(26)	$(F, C^*, V, W, 3, 6)$

(6) $(A^*, D^*, 0, 1, 8, 9)$
(7) $(B^*, E^*, 0, 2, 7, 9)$
(8) $(C^*, F^*, 0, 3, 7, 8)$

(9) $(A^*, D^*, R, U, 1, 4, 5, 8)$
(10) $(B^*, E^*, S, V, 2, 5, 6, 9)$
(11) $(C^*, F^*, T, W, 3, 4, 6, 7)$

(12) $(A^*, D^*, T, W, 1, 4, 6, 9)$
(13) $(B^*, E^*, R, U, 2, 4, 5, 7)$
(14) $(C^*, F^*, S, V, 3, 5, 6, 8)$

(15) $(A, C^*, E^*, H, W, R, 4, 7)$
(16) $(B, D^*, F^*, G, R, S, 5, 8)$
(17) $(C, E^*, A^*, H, S, T, 6, 9)$
(18) $(D, B^*, F^*, G, T, U, 4, 7)$
(19) $(E, A^*, C^*, H, U, V, 5, 8)$
(20) $(F, B^*, D^*, G, V, W, 6, 9)$

(27) (A^*, C, E, H, T, U)
(28) (B^*, D, F, G, U, V)
(29) (C^*, E, A, H, V, W)
(30) (D^*, F, B, G, W, R)
(31) (E^*, A, C, H, R, S)
(32) (F^*, B, D, G, S, T)

(33) $(B, E, C^*, E^*, W, R, 0, 2, 3, 7)$
(34) $(C, A, D^*, F^*, R, S, 0, 1, 3, 8)$
(35) $(D, B, E^*, A^*, S, T, 0, 1, 2, 9)$
(36) $(E, C, F^*, B^*, T, U, 0, 2, 3, 7)$
(37) $(F, D, A^*, C^*, U, V, 0, 1, 3, 8)$
(38) $(A, E, B^*, D^*, V, W, 0, 1, 2, 9)$

The contingent options have been arranged in compact "color matrices"

XACE	*AH*	*BG*	*CH*	*DG*	*EH*	*FG*	10 20
4*DUT*	1*A**	2*B**	3*C**	1*D**	2*E**	3*F**	47 58
6*VFW*	*RE**	*SF**	*TA**	*UB**	*VC**	*WD**	
5*SRB*	*D**9	*E**7	*F**8	*A**9	*B**7	*C**8	

which will be indicated by enclosing the scheme found in the upper left corner in quotes, thus "*X*," "*A*," etc. We are now in a position to seek systematically all the solutions of this combinatorial problem.

Case 1. Excluding all "star" schemes from one side let $(A^*B^*C^*D^*E^*F^*)^4$. Then "*R*," "*S*," · · · "*W*" imply $(RSTUVW789)^4$. Primary (6) yields the option $(0, 1)^1$ but $(0)^1(7)^4$ yields $(1)^1$ by "1"; therefore $(0,1)^1$ gives $(1)^1$. Similarly, primaries (7) and (8) give $(2)^1$ and $(3)^1$. We already have assigned three schemes to the inside, namely $(123)^1$. Now $(123)^1$(stars)^4 via "*A*," "*B*," etc. gives $(ABCDEF)^1$, and $(A)^1(WR)^4$ via "*X*" gives $(X)^1$. Primary (1) with $(ACEX)^1$ gives $(H)^2$, similarly primary (2) gives $(G)^2$. Now $(H)^2(A)^3$ via "*A*" gives $(A^*)^2$; by symmetry this applies to all the stars, thus $(A^*B^*C^*D^*E^*F^*)^2$. Primary (3) with $(ADX1)^1$ gives $(4)^1$; similarly (4) and (5) give $(5)^2$ and $(6)^2$. Now $(4)^2(1)^1$ via "1" gives

$(7)^2(0)^1$; similarly $(5)^2$ and $(6)^2$ give $(8)^2$ and $(9)^2$. We already have assigned $ABCDEFX0123$ to the inside and $A^*B^*C^*D^*E^*F^*GH456789$ to the outside. Finally "X" with $(4)^2$ provides the outside option $(RU, TW)^2$ while $(5)^2$ and $(6)^2$ provide the outside options $(RU, SV)^2$ and $(SV, TW)^2$. These compel two pairs outside, for example RU and TW. All requirements are fulfilled so S and V may be assigned in any of four ways $(SV)^2$ or $(SV)^5$ or S^2V^5 or S^5V^2. This is solution #1, the only possibility under Case 1.

Case 2. Excluding five star schemes from one side, let $(A^*B^*C^*D^*E^*)^4$. Excluding Case 1 [already considered] we have $(F^*)^1$. From "R," "U," "V," and "W" we have $(789RUVW)^4$. Now $(F^*)^1(78)^4$ via "S" and "T" yields $(S)^1$ and $(T)^1$ respectively. Primary (6) provides the option $(0,1)^1$ which combines with $(7)^4$ via "1" to give $(1)^1$. Similarly primary (7) gives $(2)^1$. Now $(1)^1(A^*D^*)^4$ gives $(A)^1$ and $(D)^1$ via "A" and "D" respectively. Likewise $(2)^1(B^*E^*)^4$ gives $(B)^1$ and $(E)^1$ via "B" and "E". Primary (32) gives $(G)^2$, which with $(BDF^*)^1$ via "B" "D" and "F" gives $(B^*D^*F)^2$. Primary (20) gives $(6)^1$, and $(6)^1(9)^4$ via "3" gives $(3)^1$; and $(3)^1(C^*)^4$ via "C" gives $(C)^1$. We now have $(CF^*ST36)^1$ contradicting primary (23). Thus Case 2 falls.

Case 3. Excluding four consecutive star schemes from one side, let $(A^*B^*C^*D^*)^4$. Excluding Case 2 [already considered] we have $(E^*F^*)^1$. From "U," "V" and "W" we have $(UVW789)^4$ and now $(E^*F^*)^1$ via "R," "S" and "T" gives $(RST)^1(7)^5$. Primary (6) provides the option $(0,1)^1$ which combines with $(7)^4$ via "1" to give $(1)^1$. Now $(1)^1(A^*D^*)^4$ via "A" and "D" gives $(AD)^1$. At this point $(BC)^2$ would be absurd. For $(BC)^2(B^*C^*)^4$ via "B" and "C" would imply $(GH23)^4$ which is incompatible with $(E^*)^1$ by virtue of "E". Accordingly $(B, C)^3$ and from the symmetry we may choose $(C)^3$.

Then primary (31) gives $(H)^2$, and then $(H)^2(ACE^*)^3$ via "A," "C" and "E" gives $(A^*C^*E)^2$. Now $(E)^2(AC)^3$ via "X" gives $(5)^2$.

Further $(C^*)^2(7)^3$ via "V" gives $(V)^2$. But then $(EA^*C^*HUV58)^4$ is incompatible with primary (19). Thus Case 3 falls.

Case 4. Let $(A^*B^*C^*)^4(E^*)^3$. Avoiding case 3 already considered we have first $(D^*F^*)^1$ and then $(A^*C^*)^2$. Thus case 4 has the symmetrical form $(D^*F^*)^1(A^*C^*)^2$. $(E^*)^3(B^*)^4$. Now $(A^*B^*C^*)^4$ via "U" and "V" gives $(UV79)^4$, while symmetry gives $(RS79)^3$, therefore $(79)^5$. Again $(D^*F^*)^1(79)^5$ via "R" and "S" gives $(RS)^1$, while symmetry gives $(UV)^2$. At this point $(H)^1$ is absurd for $(H)^1(A^*C^*)^2$ via "A" and "C" would

give $(AC)^1$, but the resulting $(E^*ACHRS)^3$ is incompatible with primary (31). This absurdity of $(H)^1$ establishes $(H)^4$, and symmetrically $(G)^3$.

Since scheme 2 cannot be on both sides, either $(2)^3$ or $(2)^4$. We remove symmetry by choosing $(2)^3$. With $(G)^3$ via "B" this gives $(BB^*)^3$. Primary (7) yields $(0)^2$, while $(0)^2(729)^3$ via "1," "2," and "3" gives $(183)^2$. Now $(13)^2(D^*F^*)^1$ via "D" and "F" compels $(DF)^2$, but the resulting $(FDA^*C^*UV0138)^2$ is incompatible with primary (37). Thus case 4 falls.

Case 5. Excluding three consecutive stars from one side, let $(A^*B^*C^*)^4$. Fallen case 3 requires $(D^*F^*)^1$, and fallen case 4 requires $(E^*)^2$. From $(A^*B^*C^*)^4$ via "U" and "V" we have $(UV79)^4$ which in turn combines with $(D^*F^*)^1(E^*)^2$ via "R" and "S" giving $(RS)^1(79)^2$.

If $(0)^4$ then primary (7) would require $(2)^1$. But this would give $(5)^1$ and $(E)^1$ via "2" and "E" respectively; and $(5E)^1(UV)^1$ via "X" would give $(BX)^1$.

But $(BEX25)^1$ is incompatible with primary (4). The absurdity of $(0)^4$ establishes $(0)^1$, and $(0)^1(79)^2$ via "1" and "3" establishes $(13)^1$ which in turn combines with $(A^*C^*)^4$ via "A" and "C" to give $(AC)^1$. Now $(ACD^*F^*RS013)^1$ and primary (34) yields $(8)^2$, and $(8)^2(0)^1$ via "2" gives $(5)^2(2)^1$. Again $(2)^1(E^*)^4$ via "E" gives $(E)^1$. But $(5)^2$ via "X" is incompatible with $(ERS)^1$. Thus case 5 falls. In any additional solutions each side must contain at least one of any three consecutive stars.

Case 6. Let $(A^*C^*E^*789)^2$.

Fallen case 5 compels $(B^*D^*F^*)^1$ which combines with $(789)^2$ via "R". . . "W" giving $(RSTUVW)^1$.

At this point $(G)^2$ is absurd. For $(G)^2(B^*D^*F^*)^1$ via "B," "D" and "F" gives $(BDF)^2$, and $(B)^2(RS)^1$ via "X" gives $(X)^2$. But $(BDFGX)^2$ is incompatible with primary (2). The absurdity of $(G)^2$ establishes $(G)^3$.

Using $(B^*D^*F^*RSTUVW)^1(G)^3$ primary options (28), (30) and (32) provide the three outside options $(D,F)^2$ and $(F,B)^2$ and $(B,D)^2$. Thus the outside contains at least two of the three schemes B, D, F and we may exploit the symmetry by choosing $(BD)^2$. Then $(BD)^2(G)^3(B^*D^*)^1$ via "B" and "D" gives $(12)^2(G)^1$. Also $(1278)^2$ via "1" and "2" gives $(045)^4$. Now $(R)^1(B5)^4$ via "X" gives $(A)^1$, and $(A)^1(1)^2$ via "A" gives $(H)^1$. In continuation $(H)^1(C^*E^*)^2$ via "C" and "E" gives $(CE)^1(3)^2$, and $(3)^2(F^*)^1$ via "F" gives $(F)^2$, restoring the symmetry, whence also $(6)^4$. Finally $(ATU)^1$ and $(ERS)^1$ and $(CVW)^1$ via "X" require $(456)^3$

which combines with $(456)^4$ to give $(456)^5$. The assignment $(0X)^2$ is completed by $(1)^2(4)^5$ via "1," and by $(B)^2(RS)^1$ via "X," respectively. The result

$$(ACEB^*D^*F^*GHRSTUVW)^1(BDFA^*C^*E^*X0123789)^2(456)^5$$

satisfies all requirements. This is solution #3.

Case 7. Excluding alternate stars from one side, let $(A^*C^*E^*)^4$. Fallen case 5 compels $(B^*D^*F^*)^1$ and $(A^*C^*E^*)^2$. From "R"$\cdots$"W" the contingencies $(7, SV)$ and $(8, TW)$ and $(9, RU)$ apply to both sides. Assigning $7 \cdot 8 \cdot 9$ to the same side is case 6, so we seek new solutions by assigning two of these schemes to one side, the third scheme to the other side. Symmetry makes $(78)^2(9)^1$ typical. We have therefore $(A^*C^*E^*78RU)^2$ $(B^*D^*F^*9STVW)^1$. Then $(R)^2(SW)^1$ via "X" implies $(45)^2$.

An inspection of "X" reveals that $(X)^3$ is absurd. For $(STX)^3$ implies $(C)^3$; $(VWX)^3$ implies $(F)^3$; $(CVW)^3$ implies $(6)^3$; and $(FF^*)^3$ via "F" implies $(3)^3$. But $(CFX36)^3$ is incompatible with primary (5): Accordingly $(X)^2$.

It is readily shown that both $(6)^2$ and $(G)^2$ are absurd. On the one hand $(6)^2(SVW)^1$ via "X" gives $(CF)^2$, and $(6)^2(9)^1$ via "3" gives $(3)^2$. But $(CFX36)^2$ is incompatible with primary (5). Therefore $(6)^3$. On the other hand $(G)^2(B^*D^*F^*)^1$ via "B," "D" and "F" gives $(BDF)^2$, but $(BDFGX)^2$ is incompatible with primary (2). Therefore $(G)^3$.

Now however primary (20) requires $(F)^2$, which combines with $(6VW)^3$ via "X" to give $(BD)^2$, and in continuation $(BDF)^2(G)^3$ via "B," "D" and "F" gives $(123)^2(G)^1$. Again $(3)^2(6)^3(9)^1$ via "3" gives $(0)^2(6)^1$. If $(H)^4$ then $(123)^2$ via "A," "C" and "E" would compel also $(ACE)^4$, but the resulting $(ACEHX)^4$ is incompatible with primary (1). Therefore $(H)^4$ is absurd, and instead we have $(H)^1$. Now $(H)^1(A^*C^*E^*)^2$ via "A," "C" and "E" gives $(ACE)^1$. The complete assignment of schemes

$$(ACEB^*D^*F^*GHSTVW69)^1(BDFA^*C^*E^*RUX01234578)^2$$

satisfies all requirements. It is solution #5.

The seven cases already investigated have served to indicate three solutions (listed as #1, #3 and #5). By virtue of case 5 and case 7 any further solutions must be obtained by assigning to each side at least one scheme from each consecutive group of three stars and at least one from each group of three alternate stars.

Case 8. Let $(B^*C^*E^*F^*)^1(A^*D^*89)^2$. The hypothesis of this case

combines with the contingencies "R"$\cdots$"W" to give $(RTUW)^1(SV7)^1$. We readily see that $(0)^2$ is absurd, for $(0)^2(7)^3$ via "1" implies $(1)^2$, but $(A^*D^*0189)^2$ is incompatible with primary (6). Therefore $(0)^3$, which combines with $(89)^2$ via "2" and "3" to give $(56)^2$. Then $(5)^2(SU)^1$ and $(6)^2(TV)^3$ via "X" imply $(BE)^2$ and $(CF)^2$ respectively. Also $(B)^2(RS)^3$ via "X" implies $(X)^2$. Again $(BEX5)^2$ and primary (4) yield $(2)^1$, while $(CFX6)^2$ and primary (5) yield $(3)^1$. Now $(BC)^2(23)^1$ via "B" and "C" yields $(GH)^2$, and $(2)^1(5)^2$ via "2" yields $(0)^1$. An inspection of "X" shows that $(BC5)^2$ infers S^4, and that $(EF6)$ infers V^4. These results combine with $(SV)^3$ to give $(SV)^5$, while $(E^*)^1(S)^5$ via "S" gives $(7)^1$.

Finally $(CEHX)^2$ and primary (1) compel $(A)^1$, while $(BFGX)^2$ and primary (2) compel $(D)^1$, so that $(A)^1(H)^2$ via "A" gives $(1)^1$, and $(D)^1(BF)^2$ via "X" yields $(4)^1$. The completed assignment

$$(ADB^*C^*E^*F^*RTUW012347)^1(BCEFA^*D^*GHX5689)^2(SV)^5$$

satisfies all requirements. This is solution #2.

Case 9. Let $(B^*C^*E^*F^*89)^1(A^*D^*)^2$. The hypothesis of this case combines with the contingencies "R"$\cdots$"W" to give $(RTUW)^2(SV7)^1$. Primaries (7) and (8) provide the outside options $(0, 2)^2$ and $(0, 3)^2$ respectively. Accordingly $(23)^2$ is compulsory, since it follows both from $(0)^3$ via these options, and from $(0)^2$ with $(89)^1$ via "2" and "3." Then $(2)^2(B^*E^*)^1$ via "B" and "E" gives $(BE)^2(GH)^1$, and $(3)^2(C^*F^*)^1$ via "C" and "F" gives $(CF)^2$. In continuation $(H)^1(A^*)^2$ via "A" gives $(A)^1(1)^2$, while $(G)^1(D^*)^2$ via "D" gives $(D)^1$. Via "X" from $(A)^1(RW)^2$ there follows $(X)^1$, and from $(D)^1(BF)^2$ follows $(4)^1$. Now $(1)^2(4)^1$ via "1" gives $(0)^2(7)^1$, and $(89)^1(0)^2$ via "2" and "3" gives $(56)^1$. Finally $(5)^1(BU)^2$ via "X" gives $(SV)^1$. The result

$$(ADB^*C^*E^*F^*GHSVX456789)^1(BCEFA^*D^*RTUW0123)^2$$

satisfies all requirements. This is solution #4.

Case 10. Excluding the special cases comprised under case 8 and case 9 let us consider $(B^*C^*E^*F^*)^1(A^*D^*)^2$. The contingencies "$R$"$\cdots$"$W$" provide the restraint $(SV7)^3$, and the options $(8, TW)$ and $(9, RU)$ applicable to both sides. Accordingly 8 and 9 must each be assigned to one side or the other. Putting $8 \cdot 9$ both outside is case 8; and putting $8 \cdot 9$ both inside is case 9. There remains only the possibility of putting one inside, the other

outside. As the two choices are symmetrically equivalent, let us consider $(8)^1(9)^2$ whence $(TW)^2(RU)^1$. Under case 10 we therefore examine

$$(B^*C^*E^*F^*RU8)^1(A^*D^*TW9)^2(SV7)^3.$$

Primary (8) reduces to the outside option $(0, 3)^2$. The trial $(0)^1$ is contrary to this option since $(0)^1(9)^2$ via "3" gives $(3)^1$. Therefore $(0)^4$, and $(8)^1(0)^4$ via "2" gives $(5)^1$. The trial E^3 is likewise absurd. For $(EE^*)^3$ via "E" gives $(2)^3$, but $(EB^*UV25)^3$ is incompatible with primary (25). The absurdity of $(E)^3$ establishes $(E)^2$, which combines with $(UV5)^3$ via "X" to give $(ACX)^2$. Now $(AA^*)^2$ via "A" compels $(H)^4$, but $(ACEHX)^4$ is incompatible with primary (1). Thus case 10 falls.

Case 11. Putting three stars on each side, [but not alternately as in case 7] let $(A^*B^*D^*)^2(C^*E^*F^*)^1$.

From "T" and "W" we obtain the option $(8, TW)$ applicable to both sides, and we may remove the symmetry by choosing $(8)^1(TW)^2$. From $(E^*F^*)^1$ in "S" follows $(S7)^3$, and from $(A^*B^*)^2$ in "U" follows $(U9)^4$. But $(B^*)^2(7)^3(C^*)^1$ in "V" gives $(V)^2(7)^1$, and $(E^*)^1(9)^4(D^*)^2$ in "R" gives $(R)^1(9)^2$.

Try $(0)^3$. Primary (8) compels $(3)^2$. Then $(3)^2(C^*F^*)^1$ via "C" and "F" gives $(CF)^2(GH)^1$, and $(GH)^1(A^*B^*)^2$ via "A" and "B" gives $(AB)^1(12)^2$. Now $(1)^2(0)^3$ via "1" gives $(4)^2$, but $(A)^1(CW4)^2$ is incompatible in "X." Then the trial $(0)^3$ fails, and instead we have $(0)^2$.

Now $(0)^2(78)^1$ via "1" and "2" gives $(12)^2(45)^1$, and $(2)^2(E^*)^1$ via "E" gives $(E)^2(H)^1$, whence $(H)^1(A^*)^2$ via "A" gives $(A)^1$. At this point primary (31) compels $(C)^2$, and $(C)^2(H)^1$ via "C" gives $(3)^2$. Then $(3)^2(F^*)^1$ via "F" gives $(F)^2(G)^1$, and $(G)^1(B^*D^*)^2$ via "B" and "D" gives $(BD)^1$.

Examining "X," we find $(W)^2(AR)^1$ gives $(6)^2$, while $(A)^1(CW)^2$ gives $(X)^1$, whence $(E)^2(XA)^1$ gives $(U)^2$, and $(5)^1(EU)^2$ gives $(S)^1$. The result $(ABDC^*E^*F^*GHRSX4578)^1(CEFA^*B^*D^*TUVW012369)^2$ satisfies all requirements. It is solution #6.

Case 12. Avoiding three stars on each side there remains but one other pattern to consider. Let $(B^*E^*)^1(A^*D^*)^2(C^*)^3(F^*)^5$. For if one side has *more* than three stars, the other side [except in case 1] must have diagonal stars, and we have the pattern $(A^*D^*)^2(B^*C^*E^*F^*)^1$ the consideration of which was concluded in case 10. If neither side has more and one side has *less* than three stars we have the pattern $(A^*D^*)^2(B^*C^*E^*)^3(F^*)^5$ in which the selection $(F^*)^5$ merely removes the symmetry. But F^*, A^*, B^* and

D^*, E^*, F^* are consecutive stars, so that to avoid duplicating case 5 we must specify $(B^*E^*)^1$ as has been done.

From $(B^*C^*E^*F^*)^3$ via "S" and "V" we have $(SV7)^3$. From $(A^*F^*)^4$ via "T" we have $(T8)^4$. From $(B^*E^*)^1(A^*D^*)^2$ via "R" and "U" we have the option $(9, RU)$ applicable to both sides.

We next show that $(2)^4$ is absurd. For $(2 \cdot 8)^4$ via "2" implies $(05)^4$ while $(B^*)^1(2)^4$ via "B" implies $(G)^1$, and $(G)^1(D^*F^*)^4$ via "D" and "F" implies $(DF)^1$. Now primary (28) compels $(U)^2$, and using the option $(9, RU)$ also $(R)^2$. Again, primary (2) reduces to $(B, X)^2$, but since $(B)^2(FS)^3$ via "X" implies $(X)^2$, therefore $(B, X)^2$ implies $(X)^2$. Then $(X)^2(DF)^1$ via "X" gives $(E)^2$. Meanwhile $(E^*)^1(2)^4$ via "E" gives $(H)^1$, and $(H)^1(A^*)^2$ via "A" gives $(A)^1$. However $(A)^1$ is incompatible with $(ERX)^2$ in "X." This absurdity establishes $(2)^1$.

Next examine $(9)^1(RU)^2$. Primary (7) yields $(0)^2$, and $(2)^1(0)^2$ via "2" gives $(5)^1(8)^2$, while $(9)^1(0)^2$ via "3" gives $(6)^1(3)^2$; and $(0)^2(7)^3$ via "1" gives $(1)^2$. Now $(3)^2(C^*)^3(F^*)^5$ via "C" and "F" gives $(CF)^2(G)^4$, and $(G)^4(1)^2$ via "D" gives $(D)^4$. With $(25)^1$ primary (4) reduces to the outside option $(B, E, X)^2$. Since $(X)^1(CDF)^4$ via "X" would give $(BE)^1$ voiding this option, therefore $(X)^4$. Then $(DRX)^4$ via "X" compels $(B)^4$. But $(BDFGX)^4$ is incompatible with primary (2). This absurdity establishes $(9)^2(RU)^1$.

Primary (6) reduces to the inside option $(0, 1)^1$. Since $(1)^3$ and $(7)^3$ via "1" yields $(0)^3$, and $(1)^2$ also yields $(0)^3$, therefore $(0)^3$.

Try $(B)^2$. Then $(B)^2(RS)^3$ via "X" gives $(X)^2$, and $(B)^2(2)^1$ via "B" gives $(G)^2$, and $(G)^2(F^*)^5$ via "F" gives $(F)^2$. Then $(FTX)^4$ via "X" gives $(D)^4$. But $(BDFGX)^4$ is incompatible with primary (2). Trial $(B)^2$ fails and therefore $(B)^3$.

In conclusion $(BRV)^3$ via "X" implies $(5)^3$, and $(BB^*)^3$ via "B" implies $(G)^3$. Then $(D^*)^2(G)^3$ via "D" gives $(1)^2$, so that the option $(0, 1)^1$ gives $(0)^1$. Now $(0)^1(1)^2$ via "1" gives $(7)^1(4)^2$, and $(0)^1(9)^2$ via "3" gives $(3)^1$. Again $(05)^3(8)^4$ via "2" gives $(8)^5$. But $(C^*F^*0378)^3$ is incompatible with primary (8). Thus case 12 falls.

All ramifications of the combinatorial problem have been investigated, and we may conclude that there are six and only six solutions, as indicated.

UNIVERSITY OF OKLAHOMA.

A NOTE ON FINITELY-ADDITIVE MEASURES.*

By ELLEN F. BUCK and R. C. BUCK

1. Introduction. In a previous paper, one of the authors has discussed a finitely-additive measure defined in I, the positive integers.[1] In Section **1**, we show that this is typical in that any separable measure in a countable space is equivalent to this measure under a 1:1 map of the space onto the integers.[2] In Section **2**, we define continuous maps of I into the interval $(0, 1)$ which are non-trivial and which generate there the usual measure. Section **3** and **4** deal with certain aspects of the topology defined by the arithmetic progressions.

2. A mapping theorem. Consider a countable class $\mathcal{D}_0$ of subsets of I, the space of positive integers. $\mathcal{D}_0$ is the minimal class containing all finite sets and arithmetic progressions, which is closed under finite union and difference. In $\mathcal{D}_0$ a finitely-additive measure Δ may be defined such that $\Delta(\{an+b\}) = 1/a$ while $\Delta(F) = 0$ if F is finite. Let $\mathcal{D}^*_0$ be the finite Carathéodory closure $\mathrm{cl}[\mathcal{D}_0, \Delta]$ of $\mathcal{D}_0$ with respect to Δ. A set S belongs to $\mathcal{D}^*_0$ and has measure $\Delta^*(S)$ if for any $\epsilon > 0$, there exist sets $B_1, B_2 \in \mathcal{D}_0$ with $B_1 \subseteq S \subseteq B_2$ and $\Delta(B_2 - B_1) < \epsilon$.

Let X be a countable space in which there is a finitely-additive measure Ω defined on a class of subsets $\mathcal{M}$, closed under finite union and difference, such that the measure of a point is zero, and $\Omega(X) = 1$. If this measure $[\mathcal{M}, \Omega]$ is in addition *separable*, so that there is a countable subclass $\mathcal{M}_0$ of $\mathcal{M}$ satisfying: (i) $\mathcal{M} = \mathrm{cl}[\mathcal{M}_0, \Omega]$, (ii) $\mathcal{M}_0$ contains no infinite sets of measure zero; and, if, (iii) $\Omega(A)$ is rational for every set of $\mathcal{M}_0$, then the following theorem holds.

THEOREM 1. *There exists a* 1:1 *map* T *of* X *onto* I *such that* $T\{\mathcal{M}_0\} \subseteq \mathcal{D}_0$, $T\{\mathcal{M}\} \subseteq \mathcal{D}^*_0$, *and if* $A \in \mathcal{M}$, *then* $\Omega(A) = \Delta^*(T(A))$.

* Received November 1, 1946.

[1] R. C. Buck, "The measure theoretic approach to density," *American Journal of Mathematics*, vol 68 (1946), pp. 560-580. This paper will be henceforth referred to as [M].

[2] This includes an unpublished result on sets having density, due, we believe, to S. Ulam.

We may clearly assume that $\mathfrak{M}_0$ contains X and all the finite subsets of X, and is closed under union and difference; for if not, we adjoin these sets and take the finite Borel extension which is still countable. Restriction (iii) is relatively unimportant. Since the measure values which occur for sets in $\mathfrak{D}^*_0$ are all the values in $[0, 1]$, the theorem, without (iii), holds if $T\{\mathfrak{M}_0\} \subseteq \mathfrak{D}_0$ is deleted. Condition (iii) is clearly necessary for the stronger form. We shall denote the cardinal number of any set S by $|S|$.

LEMMA. *Let $A \in \mathfrak{M}_0$, $B \in \mathfrak{D}_0$ be sets such that $|A| = |B|$, $\Omega(A) = \Delta(B)$. If $A_0 \subseteq A$, $A_0 \in \mathfrak{M}_0$, then there exists a set $B_0 \subseteq B$ belonging to $\mathfrak{D}_0$ such that:* (a) $\Omega(A_0) = \Delta(B_0)$, (b) $|A_0| = |B_0|$, (c) $|A - A_0| = |B - B_0|$.

Proof. (i) Suppose that $0 < \Omega(A_0) = p/q < \Omega(A) = r/s$. Write $B = \bigcup_{i=1}^{qr} P_i \cup F$ where P_i are disjoint progressions with common difference qs, and F is finite. Choose ps of the P_i in order of their first terms and call their union B_0. We see that $B_0 \subseteq B$, $B_0 \in \mathfrak{D}_0$, and $\Delta(B_0) = p/q = \Omega(A_0)$. Moreover, B_0 is infinite and $B - B_0$ consists of $qr - ps > 0$ progressions P_i and is also infinite. But so were A_0 and $A - A_0$. Thus (b) and (c) hold.

(ii) Suppose $\Omega(A_0) = 0$. Then A_0 is finite and $|A_0| + |A - A_0| = |A|$. Choose B_0 as the first $|A_0|$ terms of B.

(iii) Suppose $\Omega(A_0) = \Omega(A)$. Choose $B - B_0$ to be the first $|A - A_0|$ terms of B.

Consider the vector $\epsilon = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)$ where ϵ_i is 0 or 1. Set $\bar{\epsilon} = \Sigma \epsilon_i$ and for any set S define $\epsilon_i S$ to be S if $\epsilon_i = 1$, and to be its complement S' if $\epsilon_i = 0$. Let $\epsilon(A_1, A_2, \cdots, A_n)$ be $\bigcap_{i=1}^{n} \epsilon_i A_i$. Thus if $\epsilon \neq \epsilon'$, $\epsilon'(A_1, \cdots, A_n) \cap \epsilon(A_1, \cdots, A_n) = 0$ and $\bigcup_{\bar{\epsilon}>0} \epsilon(A_1, \cdots, A_n) = \bigcup_{1}^{n} A_i$.

Order the sets of $\mathfrak{M}_0$ into a sequence $\{A_i\}$, $(i = 1, 2, \cdots)$. Since $A_1 \subseteq X$, $\Omega(X) = 1 = \Delta(I)$ and $|X| = |I|$, we can choose, by the lemma, a set $B_1 \in \mathfrak{D}_0$ such that $\Delta(B_1) = \Omega(A_1)$ and $|B_1| = |A_1|$, $|X - A_1| = |I - B_1|$. There exists a 1:1 map T of A_1 onto B_1.

We extend this map to $A_1 \cup A_2$. Consider $A_2 - A_1 \in \mathfrak{M}_0$; it is contained in $X - A_1 \in \mathfrak{M}_0$ and, applying the lemma, there exists a set $E_1 \in \mathfrak{D}_0$, $E_1 \subseteq I - B_1$, such that $|A_2 - A_1| = |E_1|$,

$$|(X - A_1) - (A_2 - A_1)| = |(I - B_1) - E_1|$$

and $\Omega(A_2 - A_1) = \Delta(E_1)$. Let τ be a 1:1 map of $A_2 - A_1$ onto E_1. Consider now $A_2 \cap A_1$; again applying the lemma, there exists a set $E_2 \in \mathfrak{D}_0$

$E_2 \subseteq B_1$ with $|E_2| = |A_2 \cap A_1|$, $|B_1 - E_2| = |A_1 - (A_1 \cap A_2)|$ and $\Delta(E_2) = \Omega(A_2 \cap A_1)$. Since T is a 1:1 map, $T(A_2 \cap A_1)$ is a subset of B_1 with the same cardinal number relations as $A_2 \cap A_1$, and therefore as E_2. Let σ be a 1:1 map of B_1 onto itself which maps $T(A_2 \cap A_1)$ onto E_2. Set $B_2 = E_1 \cup E_2$.

Define the extended T to be τ on $A_2 - A_1$, and σT on A_1. T is now defined on $A_1 \cup A_2$. It may be easily verified that T has the following properties for $n = 2$.

(α) $T(A_i) = B_i$, $|A_i| = |B_i|$, $\Omega(A_i) = \Delta(B_i)$, $(i = 1, 2, \cdots, n)$.

(β) $T(\epsilon(A_1, A_2, \cdots, A_n)) = \epsilon(B_1, B_2, \cdots, B_n)$, all $\bar{\epsilon} > 0$.

(γ) $|\epsilon(A_1, A_2, \cdots, A_n)| = |\epsilon(B_1, B_2, \cdots, B_n)|$
$\Omega(\epsilon(A_1, A_2, \cdots, A_n)) = \Delta(\epsilon(B_1, B_2, \cdots, B_n))$, all ϵ.

Suppose that T has been defined on $S_k = \bigcup_{i=1}^{k} A_i$ with properties (α), (β), and (γ) holding for $n = k$. We extend T to $S_k \cup A_{k+1} = S_{k+1}$. Let $F_0 = A_{k+1} - S_k$. On this set T has not yet been defined. (β) and (γ) imply that $T(S_k) = \bigcup_{i=1}^{k} B_i$, $|S_k| = |\bigcup_{1}^{k} B_i|$, $\Omega(S_k) = \Delta(\bigcup_{1}^{k} B_i)$, and $|X - S_k| = |I - \bigcup_{1}^{k} B_i|$. Thus, by the lemma, there exists a set $E_0 \subseteq I - \bigcup B_i$ such that $E_0 \in \mathfrak{D}_0$, $|E_0| = |F_0|$, $|(X - S_k) - F_0| = |(I - \bigcup B_i) - E_0|$ and $\Omega(F_0) = \Delta(E_0)$. Let τ be a 1:1 map of F_0 onto E_0.

Let $F_\epsilon = A_{k+1} \cap \epsilon(A_1, A_2, \cdots, A_k)$, $\bar{\epsilon} > 0$. $T(F_\epsilon)$ has already been defined; however, by (γ) and the lemma, there is a set E_ϵ in $\epsilon(B_1, \cdots, B_k)$ such that $\Delta(E_\epsilon) = \Omega(F_\epsilon)$, $|E_\epsilon| = |F_\epsilon| = |T(F_\epsilon)|$, and

$$|\epsilon(B_1, \cdots, B_k) - E_\epsilon| = |\epsilon(A_1, \cdots, A_k) - F_\epsilon| = |\epsilon(B_1, \cdots, B_k) - T(F_\epsilon)|.$$

Let σ_ϵ be a 1:1 map of $\epsilon(B_1, \cdots, B_k)$ onto itself such that $T(F_\epsilon)$ is mapped onto E_ϵ. Set $B_{k+1} = E_0 \cup \bigcup_{\bar{\epsilon} > 0} E_\epsilon$ and define T^*, the extended T, on S_{k+1} to be τ on $A_{k+1} - S_k$ and $\sigma_\epsilon T$ on $\epsilon(A_1, \cdots, A_k)$. T^* has properties (α), (β) and (γ) for $n = k + 1$. We give here only the proof of (α); the others may be proved in a similar fashion. Suppose that $i \leq k$:

$$T^*(A_i) = T^*(\bigcup_{\epsilon_i=1} \epsilon(A_1, \cdots, A_k)) = \bigcup_{\epsilon_i=1} \sigma_\epsilon T(\epsilon(A_1, \cdots, A_k))$$

$$= \bigcup_{\epsilon_i=1} \sigma_\epsilon(\epsilon(B_1, \cdots, B_k)) = \bigcup_{\epsilon_i=1} \epsilon(B_1, \cdots, B_k) = B_i.$$

Thus, none of the sets B_i, $i \leq k$, have been altered and $|A_i| = |B_i|$, $\Omega(A_i) = \Delta(B_i)$. If $i = k + 1$, then

$$T^*(A_{k+1}) = T^*((A_{k+1} - S_k) \cup (A_{k+1} \cap S_k))$$
$$= \tau(F_0) \cup T^*(A_{k+1} \cap \bigcup_{\epsilon>0} \epsilon(A_1, \cdots, A_k))$$
$$= E_0 \cup \bigcup_{\epsilon>0} \sigma_\epsilon T(F_\epsilon) = E_0 \cup \bigcup_{\epsilon>0} E_\epsilon = B_{k+1}.$$

Since T^* is $1:1$, $|A_{k+1}| = |B_{k+1}|$. Finally,

$$\Delta(T^*(A_{k+1})) = \Delta(E_0 \cup \bigcup_{\epsilon>0} E_\epsilon) = \Delta(E_0) + \sum_{\epsilon>0} \Delta(E_\epsilon)$$
$$= \Omega(F_0) + \sum_{\epsilon>0} \Omega(F_\epsilon) = \Omega(A_{k+1}).$$

We have thus defined the mapping T recursively; since every point p of X occurs in a unit class $\{p\}$ in $\mathfrak{M}_0$, $T(p)$ is determined and its image will not be altered in further stages of the induction. Moreover, since X occurs in $\mathfrak{M}_0$, X is mapped *onto* I by T, and (α), (β), (γ) hold for all n, proving the first part of the theorem.

Now, suppose that $R \in \mathfrak{M}$; since $\mathfrak{M}$ is the finite Carathéodory extension of $\mathfrak{M}_0$, for any $\epsilon > 0$, we can find sets A_i and A_j of $\mathfrak{M}_0$ such that $A_i \subset R \subset A_j$, and $\Omega(A_i - A_j) < \epsilon$. Let $T(R) = S$. Then $T(A_i) = B_i \subset S \subset B_j$ and $\Delta(B_i - B_j) < \epsilon$ so that $S \in \mathfrak{D}^*_0$, and since $|\Omega(R) - \Delta^*(S)| < \epsilon$ for any $\epsilon > 0$, $\Omega(R) = \Delta^*(S)$. This completes the proof of the theorem.

More generally, let $[\mathfrak{M}, \Omega]$ be any finitely-additive measure on a space X, and let ϕ be a $1:1$ map of X onto X. Let $\mathfrak{M}^* = \phi\{\mathfrak{M}\}$, and define $\Omega^*(A)$, for $A \in \mathfrak{M}^*$, to be $\Omega\phi^{-1}(A)$. Then, $[\mathfrak{M}^*, \Omega^*]$ is also a finitely-additive measure on X. We call any measure obtained in this fashion a replica of $[\mathfrak{M}, \Omega]$. It is clear that any replica of a separable measure is separable.

COROLLARY. *Any separable measure on I is a contraction of a replica of $[\mathfrak{D}^*_0, \Delta^*]$.*

For, given $[\mathfrak{M}, \Omega]$ on I, there is a map T of I onto I such that $T\{\mathfrak{M}\} \subseteq \mathfrak{D}^*_0$, and $\Omega(A) = \Delta^* T(A)$. If $[\mathfrak{M}^*, \Omega^*]$ is the replica of $[\mathfrak{D}^*_0, \Delta^*]$ given by $\phi = T^{-1}$, then $\mathfrak{M} \subseteq \mathfrak{M}^*$ and $\Omega^*(A) = \Omega(A)$ for $A \in \mathfrak{M}$.

3. Continuous functions. Let Y and Z be topological spaces and f be a continuous mapping from Y into Z. Then if a measure m is defined on open sets of Y, a measure m^* can be defined for open sets O of Z by $m^*(O) = mf^{-1}(O)$. It is clearly sufficient that m need be defined only on open sets which arise as inverse images of open sets of Z. We shall discuss this construction when $Y = I$, $Z = (0, 1)$, the open unit interval, and $m = \Delta^*$; we

shall show that for a simply defined function f, m^* coincides with Euclidean length for open intervals (α, β).

We introduce a natural topology into I by choosing for neighborhoods all progressions. Thus a general neighborhood of a point $p \in I$ will be the progression $\{\lambda n + p\}$, $(n = 0, 1, 2, \cdots)$. These sets are both open and closed and so the space is totally disconnected.[3] We may define many continuous real-valued functions. For example, the characteristic function of any neighborhood is continuous. Such a function is, in a sense, a trivially continuous function even though it is not constant. For our purpose, we construct a non-trivial 1:1 real-valued continuous function.

For any $N \in I$, we may express N uniquely in the form $N = \alpha_0 + \alpha_1 r + \alpha_2 r^2 + \cdots + \alpha_n r^n$ where r and α_i are integers, $r \geq 2$ and $0 \leq \alpha_i < r$. Let $f_r(N) = \alpha_0 r^{-1} + \alpha_1 r^{-2} + \cdots + \alpha_n r^{-n-1}$. This is uniquely defined for all N and maps I in a 1:1 manner onto the set of rational numbers in $(0, 1)$ of the form A/r^k.

THEOREM 2. *The function f_r is continuous, and if O is the open interval (α, β) on $(0, 1)$, then $f_r^{-1}(O) \in \mathcal{D}^*_0$ and $\Delta^* f_r^{-1}(O) = \beta - \alpha$.*

Proof. If $f_r(N) \in (Ar^{-k}, (A+1)r^{-k})$, then $f_r(N) = Ar^{-k} + \alpha_k r^{-k-1} + \alpha_{k+1} r^{-k-2} + \cdots$ where the α_i are ultimately zero but at least one is not zero. Write $Ar^{-k} = \alpha_0 r^{-1} + \alpha_1 r^{-2} + \cdots + \alpha_{k-1} r^{-k}$; then

$$N = (\alpha_0 + \alpha_1 r + \cdots + \alpha_{k-1} r^{k-1}) + r^k(\alpha_k + \alpha_{k+1} r + \cdots)$$

and $N = a + r^k n$ where n is some positive integer. As $f_r(N)$ ranges over $(Ar^{-k}, (A+1)r^{-k})$, n ranges over all positive integers and thus $f_r^{-1}\{(Ar^{-k}, (A+1)r^{-k})\}$ is the progression $\{r^k n + a + r^k\}$ $(n = 0, 1, 2, \cdots)$, which is an open set. Similarly, the inverse image of the interval $[Ar^{-k}, (A+1)r^{-k})$ is the progression $\{r^k n + a\}$ $(n = 0, 1, 2, \cdots)$. Thus the inverse image of (Ar^{-k}, Br^{-k}) is the sum of $B - A$ *disjoint* progressions of the form $\{r^k n + a\}$, is open and belongs to $\mathcal{D}_0$. Hence f_r is continuous. Moreover, $\Delta f_r^{-1}\{(Ar^{-k}, Br^{-k})\} = r^{-k}(B - A)$.

For any (α, β) in $(0, 1)$, we may choose rational numbers such that for large k,

---

[3] This implies that the space is metrizable; for example, let

$$\rho(x, y) = \sum_{a,b=1}^{\infty} |f_{a,b}(x) - f_{a,b}(y)| \, 2^{-a-b}$$

where $f_{a,b}$ is the characteristic function of $\{an + b\}$. The space is not compact. Topologies of this type on the additive group of integers are often used as illustrations—e.g., Pontrjagin, *Topological Groups*, p. 57.

$$A_k r^{-k} \leq \alpha \leq A'_k r^{-k} < B'_k r^{-k} \leq \beta \leq B_k r^{-k}$$

and

$$\lim A_k r^{-k} = \lim A'_k r^{-k} = \alpha, \qquad \lim B_k r^{-k} = \lim B'_k r^{-k} = \beta.$$

If $S = f_r^{-1}\{(\alpha, \beta)\}$, the sets of $\mathfrak{D}_0$ corresponding to $(A'_k r^{-k}, B'_k r^{-k})$ and $(A_k r^{-k}, B_k r^{-k})$ approximate S from within and without such that the Δ-measure of their difference is $(B_k - A_k) r^{-k} - (B'_k - A'_k) r^{-k}$ which approaches zero. Thus the set S belongs to $\mathfrak{D}^*_0$ and $\Delta^*(S) = \lim (B_k - A_k) r^{-k}$ or $\Delta^*(S) = \beta - \alpha$.

4. Decompositions. We cannot conclude that the inverse image of an arbitrary open set in $(0, 1)$ is measurable Δ^* for there are open sets in I which do not belong to $\mathfrak{D}^*_0$. Any open set in I is a union of progressions; since the union and difference of any two progressions is either null or a finite union of disjoint progressions, we may express any open set as a countable union of *disjoint* progressions. Let $\mathfrak{D} \supset \mathfrak{D}^*_0$ be the class of subsets of I for which a density exists, defined as $D(S) = \lim S(N)/N$ where $S(N)$ is the number of terms of S not greater than N.

THEOREM 3. *Let S be an open set in I, $S = \bigcup B_i$, where $B_i = \{b_i n + a_i\}$ and $B_i \cap B_j = 0$ if $i \neq j$. Let $A = \{a_i\}$. Then, $S - A \in \mathfrak{D}$ and $D(S - A) = \beta = \Sigma\Delta(B_i)$.*

Proof. Clearly $S(N) = \Sigma B_i(N)$. Since $B_i = \{b_i n + a_i\}$,

$$B_i(N) = \begin{cases} 1 + [(N - a_i)/b_i] & \text{if } a_i \leq N \\ 0 & \text{if } a_i > N. \end{cases}$$

Thus

$$S(N) = \sum_{a_i \leq N} 1 + \sum_{a_i \leq N} [(N - a_i)/b_i]$$

or

$$(S(N) - A(N))/N = 1/N \sum_{a_i \leq N} [(N - a_i)/b_i] = u_N.$$

Since

$$u_N \leq 1/N \sum_{a_i \leq N} N/b_i = \sum_{a_i \leq N} 1/b_i \leq \beta,$$

$\overline{\lim}\, u_N \leq \beta$. For the other direction, choose some index k; if N is sufficiently large so that $N > a_k$,

$$u_N \geq \sum_{a_i \leq a_k} [(N - a_i)/b_i]/N$$

and as N increases,

$\underline{\lim}\, u_N \geq \sum_{a_i \leq a_k} 1/b_i$. Then letting k increase, $\underline{\lim}\, u_N \geq \beta$.

COROLLARY 1. *$S \in \mathfrak{D}$ if and only if $A \in \mathfrak{D}$ and then $D(S) = D(A) + \beta$.*

Since I is open, we can express it as the union of disjoint progressions; in particular, $I = \bigcup_{\lambda=1}^{\infty} 2^\lambda n + 2^{\lambda-1}$. More generally, if r is any integer, $r \geq 2$, $I = \bigcup B_i$ where $B_i = \{r^i n + a_i\}$, and $B_i \cap B_j = 0$ if $i \neq j$. To effect this we need only choose $a_1 = 1$, and choose $a_{\lambda+1}$ as the least integer in $I - \bigcup_{i=1}^{\lambda} B_i$. From this decomposition we see at once that there are a non-countable number of sets both open and closed. Observe also that

$$\beta = \sum \Delta(B_i) = \sum 1/r^i = 1/(r-1)$$

which may be made arbitrarily small. Applying the above theorem to such a decomposition, we obtain in particular:

COROLLARY 2. *There exist open sets not in* $\mathfrak{D}$.

Let $I = \bigcup \{3^i n + a_i\} = \bigcup B_i$, $B_i \cap B_j = 0$. Since $I \in \mathfrak{D}$, $A = \{a_i\}$ belongs to $\mathfrak{D}$ and $D(A) = 1 - \beta = 1/2 > 0$. Select a subset A_0 of A which does not have a density, and define S as $\bigcup_{a_i \in A_0} B_i$. By Corollary 1, this open set cannot belong to $\mathfrak{D}$.

Such decompositions of I also yield open sets which belong to $\mathfrak{D}$ but not to $\mathfrak{D}^*_0$. A set E is everywhere dense if its closure is I; then, every progression meets E in an infinite number of points.[4] This is equivalent to $\bar{\Delta}(E) = 1$ where $\bar{\Delta}(E) = \inf \Delta(B)$ for $E \subseteq B$ and B in $\mathfrak{D}_0$. [M].

THEOREM 4. *If* $I = \bigcup \{r^i n + a_i\}$ *is a disjoint decomposition of* I, $A = \{a_i\}$, *then* $A' = I - A$ *is an open everywhere dense set belonging to* $\mathfrak{D}$, *but if* $r > 2$, *not to* $\mathfrak{D}^*_0$.

Proof. $A' = \bigcup \{r^i n + a_i + r^i\}$ and is open but not closed. To show that A' is everywhere dense it is sufficient to show that every progression $\{\lambda m + a_j\}$ meets A' infinitely often for every λ and j $(\lambda, j = 1, 2, \cdots)$. For then the closure of A' is I. If two progressions have one point in common, their intersection is infinite. If we choose $i = j$, $m = r^j$, $n + 1 = \lambda \geq 1$, then for any λ and j, $\lambda m + a_j = r^i n + a_i + r^i$; thus $\bar{\Delta}(A') = 1$. By the previous theorem, A' belongs to $\mathfrak{D}$ and $D(A) = (r-2)/(r-1)$. Since D agrees with Δ on $\mathfrak{D}_0$, $\bar{\Delta}(A) \geq D(A)$. If $C_k = \bigcup_{i=1}^{k} \{r^i n + a_i + r^i\}$, then $A \subset C'_k$ and

[4] Added in proof. Everywhere dense sets predominate. A measure may be defined in 2^I by a dyadic mapping of this onto $(0, 1)$. [M] By a standard application of Rademacher functions, one may show the following: *if* $A \subset I$ *does not have unit lower density, the class of subsets of* A *has zero measure.* Hence, almost every subset of I is everywhere dense in I (extremal) and $\mathfrak{D}_0{}^*$ has measure zero. This answers a question raised in [M, p. 580].

$$\bar{\Delta}(A) \leq \lim \Delta(C'_k) = 1 - \sum_1^\infty 1/r^4.$$

Hence $\bar{\Delta}(A) = (r-2)/(r-1)$. If $r > 2$, $\bar{\Delta}(A) + \bar{\Delta}(A') > 1$ and A' is non-measurable. We remark that A is nowhere dense in I but $D(A)$ may be arbitrarily close to $D(I) = 1$.

5. Special sets. Since the topology is inherently arithmetic, it is natural to look for connections with the theory of numbers.

THEOREM 5. *The set of primes, P, is nowhere dense, has 1 as its only limit point but is not convergent.*

Since $\Delta^*(P) = 0$, the set of composites $C = I - P - \{1\}$ is everywhere dense. [M]. But $C = \bigcup_{p \in P} \{np + p\}$ and so C is open, $P \cup \{1\}$ is closed and therefore nowhere dense. P has 1 as a limit point if and only if every progression $\{\lambda n + 1\}$ contains an infinite number of primes; this follows from Dirichlet's theorem.[5] For any $q \neq 1$, the neighborhood $\{qn + q\}$ contains at most one prime and q is not a limit point of P. Since $\{4n + 3\}$ contains an infinite number of primes but does not contain 1, P does not converge. We can, of course, choose a subset of the primes convergent to 1. For example, let p_k be the least prime in $\{(k+1)!\, n + 1\}$ distinct from $p_1, p_2, \cdots, p_{k-1}$.

THEOREM 6. *The set $\{m^2\}$ is closed, perfect, and nowhere dense.*

Let q be a limit point of $\{m^2\}$. Then for any λ, $\lambda n + q$ must belong to $\{m^2\}$ for an infinite number of n. Thus, $x^2 - q \equiv 0 \pmod{\lambda}$ has a solution for each λ. It then follows that q is a square. If $p = \alpha^2$, then for each λ, and for all t, $(\alpha + \lambda t)^2 = \alpha^2 + \lambda(t^2\lambda + 2t\alpha)$. Hence $\{\alpha^2 + \lambda n\}$ intersects $\{m^2\}$ infinitely often. Thus $\{m^2\}$ is closed and perfect. Moreover, $\Delta^*(\{m^2\}) = 0$ and so $\{m^2\}$ is nowhere dense. [M].

THEOREM 7. *$S = \{[\alpha n + \beta]\}$ where α is larger than 1 and irrational, $\beta \geq 0$, is neither open nor closed, and is everywhere dense.*

The proof follows immediately since $\bar{\Delta}(S) = 1 = \bar{\Delta}(S')$. [M]. This result is dual to the classical one which states that if α is irrational, the fractional parts of $\alpha n + \beta$ are everywhere dense in $(0, 1)$.

WELLESLEY COLLEGE AND
SOCIETY OF FELLOWS, HARVARD UNIVERSITY.

---

[5] Landau, *Vorlesungen uber Zahlentheorie*, vol. 1. An elementary proof of the case used here can also be given. G. D. Birkhoff and H. S. Vandiver, "On the integral divisors of $a^n - b^n$," *Annals of Mathematics*, Ser. 2, vol. 5 (1903), pp. 173-180.